# Channel access optimization with adaptive congestion pricing for cognitive vehicular networks: an evolutionary game approach

Article (Accepted Version)

# Channel Access Optimization with Adaptive Congestion Pricing for Cognitive Vehicular Networks: An Evolutionary Game Approach

Daxin Tian, *Senior Member, IEEE,* Jianshan Zhou, Yunpeng Wang, Zhengguo Sheng, Xuting Duan, and Victor C.M. Leung, *Fellow, IEEE*

**Abstract**—Cognitive radio-enabled vehicular nodes as unlicensed users can competitively and opportunistically access the radio spectrum provided by a licensed provider and simultaneously use a dedicated channel for vehicular communications. In such cognitive vehicular networks, channel access optimization plays a key role in making the most of the spectrum resources. In this paper, we present the competition among self-interest-driven vehicular nodes as an evolutionary game and study fundamental properties of the Nash equilibrium and the evolutionary stability. To deal with the inefficiency of the Nash equilibrium, we design a delayed pricing mechanism and propose a discretized replicator dynamics with this pricing mechanism. The strategy adaptation and the channel pricing can be performed in an asynchronous manner, such that vehicular users can obtain the knowledge of the channel prices prior to actually making access decisions. We prove that the Nash equilibrium of the proposed evolutionary dynamics is evolutionary stable and coincides with the social optimum. Besides, performance comparison is also carried out in different environments to demonstrate the effectiveness and advantages of our method over the distributed multi-agent reinforcement learning scheme in current literature in terms of the system convergence, stability and adaptability.

**Index Terms**—Cognitive vehicular networks, vehicle-to-infrastructure (V2I) communications, evolutionary game theory, opportunistic spectrum access, dynamic pricing.

## 1 Introduction

Vehicular networks build the fundamental blocks of connected vehicles and need to serve rapidly increasing requirements of future intelligent transportation systems (ITS) and tremendous vehicular telematic applications, such as bandwidth-intensive real-time traffic monitoring, network-wide traffic big data collection, large-volume vehicular content transmissions, and some others. Considerable research efforts have been devoted to the field of vehicular networks. For instance, U.S. FCC has allocated an exclusive-use frequency band of 75 MHz for dedicated short range communication (DSRC) in vehicular environments, which is between 5.850 GHz and 5.925 GHz [1]. However, due to the inherent limitations, none of DSRC and mobile cellular networks can independently and fully address all types of growing applications in terms of network coverage, bandwidth as well as quality of service (QoS). In this context, vehicular hybrid channel access paradigms based on cognitive radio (CR) [2], [3] have been developed to enhance the connectivity of vehicles and improve the utilization of constrained spectrum resources.

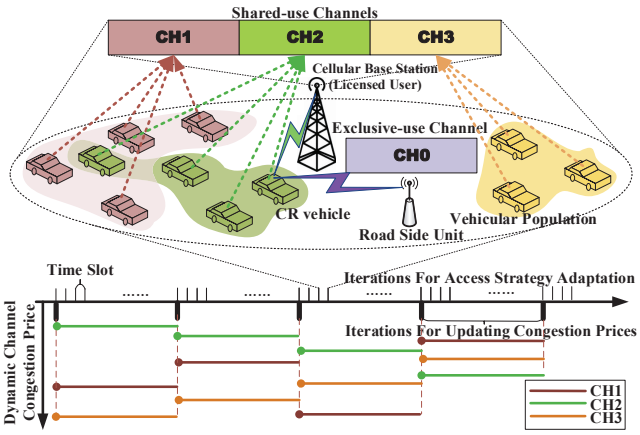



Fig. 1. An explanatory scenario of hybrid channel access in a cognitive vehicular network and the timing diagram of access strategy evolution of vehicular populations and dynamic channel pricing.

We consider a cognitive vehicular network as shown in Fig. 1, where vehicular nodes with CR can use multi-mode radio interfaces deployed onboard to dynamically access two types of radio channels, one of which is the vehicular exclusive channel, such as the DSRC channel, that is particularly reserved for reliability-critical vehicular communications, and the other is known as the shared-use channel [4], [5], which is provided by a licensed user (e.g., a cellular base station). In the same service cell of the licensed user, a great number of vehicular nodes as unlicensed users need to make decisions on accessing the opportunistic


- *D. Tian, J. Zhou, Y. Wang and Xuting Duan are with Beijing Advanced Innovation Center for Big Data and Brain Computing, Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems & Safety Control, School of Transportation Science and Engineering, Beihang University, Beijing 100191, China (e-mail: dtian@buaa.edu.cn, jianshanzhou@foxmail.com)*
- *Z. Sheng is with Department of Engineering and Design, the University of Sussex, Richmond 3A09, UK.(E-mail: z.sheng@sussex.ac.uk)*
- *V. Leung is with Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, B.C., V6T 1Z4 Canada.(E-mail: vleung@ece.ubc.ca)*

shared-use channel. The integration of both the exclusive-use and the shared-use channels for CR-enabled vehicles can be regarded as a cooperation between heterogeneous spectrum resources. In such a cognitive vehicular network, two fundamental questions arise:

- *how to model the radio channel access by a large population of competitive vehicular nodes and analyze their dynamic behaviors*?
- *what is an incentive to induce vehicular nodes to achieve a socially optimal and balanced utilization of the spectrum resources, if vehicles make access decisions in a self-interest-driven manner*?

In reality, individual users are not usually with social responsibility. Indeed, it is crucial and challenging to understand the dynamics of hybrid channel access of large-scale vehicular users with bounded rationality in pursing their benefit. The questions mentioned above become much more challenging when the vehicular populations are changing over time (i.e., the number of vehicular users within the service area is time-varying). In this situation, a dynamically adaptive incentive mechanism is needed, such that the vehicular nodes are able to adapt their access strategies efficiently to the changing environment. Then, another basic question follows:

- *How well can an incentive mechanism applied to selfish users perform in terms of the adaptability and the optimality of the resulting system*?

Our work considers the fundamental questions above. To be specific, we pose the key problem aforementioned as a noncooperative game. Motivated by evolutionary dynamics of biological populations in nature where self-interest-driven individuals usually replicate the behavior of others who can lead to higher benefit [6], [7], we develop a dynamic evolutionary game-theoretic framework. Moreover, different from much existing literature in which the Nash equilibrium is usually considered as the system solution, we reveal the inefficiency of the Nash equilibrium in the sense that the total cost utility of the vehicular nodes is not minimized. That is, because the Nash equilibrium in a noncooperative game are generally inefficient as widely reported in existing literature [8], [9], [10], the Nash equilibrium is not necessarily the optimal point and may not be suitable to be the final solution. We devise a proper incentive mechanism as the dynamic control of channel access behavior of selfish vehicular nodes, which is based on dynamic congestion pricing. Specifically, vehicular nodes are required to pay a certain price to the spectrum resource provider according to their selected channel, and such a price is designed according to the time-varying congestion degree associated with a channel. We formulate the price of a shared-use channel based on the concept of marginal social cost that is originally proposed by Pigou [11]. Accordingly, a vehicular node should be charged by the marginal social cost increased by its participation in channel competition. It is also worth pointing out that the dynamic pricing based the marginal social cost has been widely adopted as an effective means of macro control in many other fields such as macroeconomics, market management, transportation, and some others [11], [12].

However, from the practical implementation perspective, channel pricing of the licensed provider and channel access decision-making of individual vehicles should not be operated on the same time scale. Directly imposing a real-time changing channel congestion price (this means that the time resolution of updating the price is as the same as that of making channel access decisions) on the competing vehicular nodes may be impractical in actual situations. This is due to three major facts: i) it is usually cumbersome and annoying for vehicular users to respond to the fast-changing price of a channel at the same moment when they are making decisions on accessing the channel; ii) vehicular nodes should first learn the channel congestion price and then actually determine which channel to access; and iii) the limited transmission capacity of the backhaul connection between vehicle nodes and resource providers could delay the computing information exchanged between them. At this point, it is more meaningful to allow vehicular nodes to get prior knowledge of the channel prices, so that vehicular nodes can make informed decisions before actually accessing a channel. Hence, we propose a delayed dynamic pricing mechanism, where the congestion price of each channel is updated on the basis of a relatively large time period, i.e., a time interval that is larger than the time slot for evolution of the cognitive vehicular network. Within the large time interval, the channel price stays constant (as shown in an explanatory scenario in Fig. 1). In this way, the access strategy adaptation and the update of the channel prices are performed in an asynchronous manner. Besides, when introducing the delayed pricing mechanism to regulate the discrete evolutionary dynamics of the vehicles' opportunistic channel access, an additional challenging question arises that whether the evolutionary dynamics can still guarantee its convergence, which is also answered in detail in this paper.

The main contributions of this work can be summarized as follows:

- We present an evolutionary game model to analyze the evolution of self-interest-driven CR-enabled vehicular nodes contending for the shared-use channels. We prove that it is an exact potential game and guarantees the existence and the stability of the Nash equilibrium.
- To further address the inherent inefficiency of the Nash equilibrium-based solution, we propose a delayed dynamic pricing mechanism as an incentive for regulating vehicular nodes' behavior based on the marginal costs of the channels. By integrating the proposed pricing mechanism with the evolutionary game, we develop a discretized replicator dynamics with the delayed congestion pricing to enable the strategy learning and adaptation of vehicular nodes and the price updating of the channels in an asynchronous manner. We further investigate the effect of the pricing delay on the proposed discretized replicator dynamics and prove that its Nash equilibrium coincides with the socially optimal solution.
- We theoretically and numerically study the asymptotic convergence and stability of the proposed evolutionary dynamics with the delayed dynamic pric-

ing using two types of learning rates, i.e., a dynamic learning rate and a constant learning rate, respectively. We analytically derive the bound of the difference between the proposed evolutionary dynamics and that without a delay in pricing, and further obtain the theoretical gap between the steady performance of the proposed evolutionary dynamics and the ideal optimum, which is shown to be in the same order of the implemented learning rate.

The remainder of this paper is organized as follows. In Section 2, we present a review of the related work. Section 3 describes the system model. In Section 4, an evolutionary game model is formulated for decision-making behavior of vehicular nodes in hybrid channel access. Subsequently, a replicator dynamics algorithm with a delayed dynamic pricing mechanism is developed in Section 5. Section 6 presents numerical results and discussions. Finally, Section 7 concludes this work as well as outlines our future work.

## 2 Related Work

The emerging CR, sometimes termed software defined radio, is also introduced to vehicular nodes, which has spawned cognitive vehicular networks, such that vehicular nodes can obtain local intelligence and self-cognition [2], [5]. The authors of [13] take into consideration shared-use and exclusive-use channels in a general cognitive vehicular network, for which constrained Markov process formulations are developed. In despite of appealing opportunities cognitive vehicular networks can provide, there exist some challenging fundamental issues in modeling, control and optimization aspects to be handled, such as spectrum sensing and sharing [3], [4]. There have been numerous research efforts made on developing different cooperative spectrum sensing solutions for accurate detection of the licensed users' channel state, e.g., [14]. Dynamic spectrum sharing or dynamic channel access is another key functionality of CR networks, which is focused on in our work.

Indeed, a vast amount of other works have also been published in the field of dynamic spectrum access, among which the game theory is widely employed as a fundamental and powerful tool for modeling of decision-making behavior of large-scale populations [15]. Different types of game-theoretic models have been proposed, such as the stochastic learning-based potential game formulations [16], the coalitional game models [17], etc.. In [18], the authors also propose an evolutionary game framework to analyze the dynamics of access network selection of users. Some other noncooperative game-theoretic formulations can be referred to in [19], [20], while another study [21] presents a cooperative game model to address the allocation of secondary spectral resources among a population of CR secondary users. There are also many research efforts that have been made based on reinforcement learning to address the issue of channel allocation such as [22], [23], [24]. In [22], the deep Q-network (DQN), a representative reinforcement learning that is originally proposed by DeepMind [25] and currently has been widely applied in many other fields such as robot control, computer vision, human-machine interactive games, etc., is adopted for dynamic multichannel access. In the application situation, only a single node is considered. In [23], a multi-agent reinforcement learning is used to make a decision on the selection of a radio access technology (RAT) from a feasible set, which can provide Nash equilibrium strategies. In [24], a distributed learning mechanism is proposed based on evolutionary game theory for spectrum access, which can well achieve an equilibrium with guaranteed evolutionary stability. The authors also show that the evolutionary dynamics-based learning mechanism can outperform the traditional distributed Q-learning mechanism. In these aforementioned works, the evolutionary equilibrium or Nash equilibrium is considered as the system solution and the potential inefficiency of the evolutionary equilibrium or Nash equilibrium has been neglected. Nonetheless, from the global perspective, Nash-equilibrium solutions may not necessarily optimize the overall benefit of the decision makers.

To optimize the spectrum sharing in CR networks, many different dynamic pricing policies have been devised as a kind of adaptive control on the decision-making behavior of the CR users. There are a number of recent works, such as [26], [27], [28], [29], that deal with the price optimization problem with a hierarchical Stackelberg game-theoretic framework. In these works, the optimal spectrum price is dependent on the Stackelberg equilibrium points. In the works [30], [31], the competition among secondary users is formulated as a noncooperative game where a pricing-based policy is also applied to adapt the users' spectrum buying decisions. In their models, and the evolutionary dynamics is adopted to iteratively adjust the users' access strategies, and the Nash equilibrium obtained is considered as the solution to the game. Some other works [32], [33] have investigated different pricing models for spectrum trading, including market-equilibrium, competitive, and cooperative pricing models. However, the issue about the inefficiency of a Nash equilibrium-based solution has not been fully considered in the above works. Differently, in [10], the authors present the interaction between primary service providers and secondary users as an oligopoly market and establish a dynamic Bertrand game, where the Nash equilibrium is revealed inefficient. Although [10] meaningfully proves the inefficiency of Nash equilibrium, they have not investigated the impact of pricing delays on the repeated game.

Here, the essential goals of our work are twofold in modeling and optimization aspects. First, we propose a dynamic evolutionary game framework with a discretized replicator dynamics to model the evolution of opportunistic channel access of vehicular nodes in a cognitive vehicular network. Then, with respect to optimization, our objective is to induce the vehicular nodes to approach the global optimal performance through the repeated game. Different from much game-theoretic approach-based literature, we theoretically and numerically show the inherent inefficiency of the Nash equilibrium, and propose a channel congestion pricing mechanism based on the marginal social cost to motivate selfish vehicular nodes to achieve a social optimum of the spectrum allocation. Moreover, we account for information delay in spectrum pricing and investigate how a delayed spectrum price can affect the convergence and steady performance of the proposed discretized evolutionary dynamics, which has been mostly ignored in the aforementioned literature on spectrum sharing and spectrum pricing. We

also study the convergence and steady performance of the proposed dynamics using a dynamic learning rate and a static learning rate, respectively, analytically deriving the order of the difference between the discretized evolutionary dynamics with the proposed delay pricing mechanism and the corresponding evolutionary dynamics with a real-time pricing under these two types of learning rates. The theoretical developments in this paper can shed light on a trade-off between the adaptability and the optimality of channel access management in cognitive vehicular networks in terms of actual system design.

## 3 System Model

### 3.1 Hybrid channel access

We consider a V2I scenario, e.g., Fig. 1, where there are a large number of CR-enabled vehicles within a service area, $\mathcal{N}$, overlaid with a licensed user (e.g., a base station). Let the total number of vehicular nodes in $\mathcal{N}$ be $N$, i.e., $N = |\mathcal{N}|$. Here, we remark that $N$ is indeed an environmental parameter which is usually time-varying due to the vehicle mobility. Nevertheless, the time scale (e.g., several ten seconds or several minutes) for changing the macro-level spatial distribution of vehicles within the area is much larger than a time slot (e.g., several milliseconds) of micro-level physical communications [13], such that the environmental parameter $N$ can be regarded as a constant within the time scale of changing the spatial distribution of vehicles. In the CR-enabled V2I communication situation, the licensed user shares a number of licensed channels with the $N$ vehicular nodes. We denote the set of the shared-use channels by $\mathcal{M}$ and let $M = |\mathcal{M}|$. We also remark that the CR-enabled vehicular network shares some characteristics (such as the population behavior in competing for limited spectrum resources, nodes' mobility) with some other types of CR-enabled systems such as cognitive radio wireless sensor networks (CR-WSNs), cognitive heterogeneous networks, etc., the methodological approach proposed in this paper can be extended and applied to address the issue of channel access optimization in those situations. Once the channel model and the physical-layer transmission are appropriately characterized based on the targeted situation, we can map the practical spectrum resource access problem in CR-enabled networks into proper evolutionary game-theoretic formulations. In the next subsection, we will elaborate on the channel and physical-layer model for the focused CR-enabled vehicular network.

In the slotted fashion, the licensed user first broadcasts the availability information on the shared channels to the vehicular nodes over a broadcast channel (i.e., a dedicated control channel). Then, a vehicle can send a request to the licensed user through the control channel when it would like to access one of the shared-use channels. After receiving the access request, the licensed user allocates a number of time slots to the vehicle. This allocation of time slots can be implemented by adopting a TDMA-based MAC mechanism such as IEEE 802.16d/e standard. In addition, we assume that each CR vehicular node has two different interfaces [4], [13], [28], such that they are able to perform data transmissions on the shared-use channels (e.g., **CH1**, **CH2** and **CH3** in Fig. 1) provided by the licensed user and a dedicated channel (e.g., **CH0** in Fig. 1), simultaneously. The exclusive channel is assumed to be reserved for short-range vehicular communications and controlled by a local provider near the tagged vehicle, for instance, a road side unit (RSU). Each vehicular node cannot access a shared-use channel and the exclusive-use channel when the selected channels are occupied. The competition occurs among multiple vehicular nodes to select the same shared-use channel. Following current literature [16], [34], we assume that each vehicular node contends to transmit on the same channel $i \in \mathcal{M}$ in each time slot with the probability $\alpha_i \in (0,1)$ and that in one time slot a vehicular node can only select one shared channel even when multiple spectrum opportunities may be available to it. It is also assumed that the vehicular nodes can accurately sense the shared channels and implement a carrier-sense multiple access (CSMA) technique to address collisions when sharing the same channel (in fact, many cooperative spectrum sensing schemes proposed in current literature, such as [14], [17], can be used to achieve accurate channel sensing by the nodes, which are out of the scope of this paper). In the competition scenario, we use $n_i \in \mathbb{Z}_0^+$ to denote the number of vehicular nodes that are contending for the shared channel $i \in \mathcal{M}$, and we have $N = \sum_{i=1}^{M} n_i$.

### 3.2 Channel and physical-layer model

We consider the Nakagami-$m$ distribution to characterize the channel fading in the cognitive vehicular network [13], [35], [36]. Let $\overline{\omega}_1$ and $\overline{\omega}_2$ be the average signal-to-noise ratio (SNR) for the shared channel and the exclusive channel, respectively. In the physical layer, the vehicular nodes are considered to adopt the adaptive modulation in order to enhance data transmissions. To be specific, we assume that the vehicular wireless transceiver can support $L$ different modulation rates on the shared channel $i$, which are denoted by $\mathcal{C}_i = \{c_{l,i}, l = 1, 2, \ldots, L\}$ where $c_{l-1,i} < c_{l,i}$ for $l = 2, 3, \ldots, L$. Each $c_{l,i}$ is associated with two SNR thresholds, $\Gamma_{l-1}$ and $\Gamma_l$ ($\Gamma_{l-1} < \Gamma_l$). When the received SNR at the receiver, $\omega$, satisfies $\omega \in [\Gamma_{l-1}, \Gamma_l)$, the onboard receiver implements the corresponding modulation rate $c_{l,i}$, such that it can transmit $c_{l,i}$ packets over the selected shared channel $i$ in a time slot. At this point, the received SNR on a channel can be divided into $L$ non-overlapping SNR intervals, i.e., $\{[\Gamma_0, \Gamma_1), [\Gamma_1, \Gamma_2), \ldots, [\Gamma_{L-1}, \Gamma_L)\}$ where we set $\Gamma_0 = 0$ and $\Gamma_L = +\infty$. Additionally, we assume that each vehicular node is able to transmit $c'_l$ packets on the exclusive channel per time slot if the received SNR on this channel falls within the $l$-th SNR interval $[\Gamma_{l-1}, \Gamma_l)$.

Let $\kappa$ be an indicator, where $\kappa = 1$ is for the case of data transmissions on the shared channel, while $\kappa = 2$ is for the exclusive channel. According to the Nakagami-$m$ distribution, we can have

$$\text{Prob}\{\omega_\kappa \le x\} = \frac{\gamma\left(m, \frac{m}{\overline{\omega}_\kappa} x\right)}{\Gamma(m)} \tag{1}$$

for $\kappa = 1, 2$, where $\Gamma(m)$ is a Gamma function characterized by the parameter $m$, $\Gamma(m) = \int_0^\infty e^{-s} s^{m-1} ds$, and $\gamma\left(m, \frac{m}{\overline{\omega}_\kappa} x\right)$ is a lower incomplete Gamma function, $\gamma\left(m, \frac{m}{\overline{\omega}_\kappa} x\right) = \int_0^{\frac{m}{\overline{\omega}_\kappa} x} e^{-s} s^{m-1} ds$.

Denote by $\mathrm{Prob}_{\mathrm{SNR}}\{\kappa, l\}$ the probability of the received SNR falling within $[\Gamma_{l-1}, \Gamma_l)$. Namely, $\mathrm{Prob}_{\mathrm{SNR}}\{\kappa, l\} = \mathrm{Prob}\{\omega_\kappa \in [\Gamma_{l-1}, \Gamma_l)\}$ can be obtained from (1) as follows

$$\mathrm{Prob}_{\mathrm{SNR}}\{\kappa, l\} = \frac{\gamma\left(m, \frac{m}{\overline{\omega}_\kappa}\Gamma_l\right) - \gamma\left(m, \frac{m}{\overline{\omega}_\kappa}\Gamma_{l-1}\right)}{\Gamma(m)} \quad (2)$$

for $\kappa = 1, 2$. With $\omega_\kappa \in [\Gamma_{l-1}, \Gamma_l)$ and (2), we can also calculate the average packet error rate (PER) by [13], [37]

$$p(\kappa, l) = \frac{1}{\mathrm{Prob}_{\mathrm{SNR}}\{\kappa, l\}} \frac{a_l}{\Gamma(m)} \left(\frac{m}{\overline{\omega}_\kappa}\right)^m \frac{\gamma(m, b_l\Gamma_l) - \gamma(m, b_l\Gamma_{l-1})}{(b_l)^m} \quad (3)$$

where $a_l$ and $b_l$ are two tunable parameters that can be evaluated by fitting the theoretical equation to the exact PER measurement [13], [37]. By using an automatic repeat request (ARQ) scheme, the probability of success in transmitting $c_{l,i}$ packets over the shared channel $i$ in mode $l$ ($1 \le l \le L$) in a time slot can be determined as

$$\mathrm{Prob}_1\{c_{l,i}, i\} = \mathrm{Prob}_{\mathrm{SNR}}\{1, l\}\left(1 - p(1, l)\right)^{c_{l,i}}. \quad (4)$$

Similarly, for the case of the exclusive channel, the probability can be expressed as

$$\mathrm{Prob}_2\{c_l'\} = \mathrm{Prob}_{\mathrm{SNR}}\{2, l\}\left(1 - p(2, l)\right)^{c_l'}. \quad (5)$$

Considering that each vehicular node can use two radio interfaces to transmit packets, the average number of packets that can be successfully transmitted by a vehicular node contending to access the shared channel $i$ is expressed by

$$r_i(n_i) = \frac{\mathrm{C}_{n_i}^1 \alpha_i (1-\alpha_i)^{n_i - 1}}{n_i} \times \left(\sum_{l=1}^{L} c_{l,i}\mathrm{Prob}_1\{c_{l,i}, i\} + \sum_{l'=1}^{L} c_{l'}'\mathrm{Prob}_2\{c_{l'}'\}\right), \quad (6)$$

where the term $\frac{1}{n_i}$ represents the probability of a vehicular node successfully contending for transmission on the channel $i$ when considering a uniformly random competition among $n_i$ vehicular nodes, and $\mathrm{C}_{n_i}^1 \alpha_i (1-\alpha_i)^{n_i - 1}$ denotes the probability of no transmissions by the other $n_i - 1$ nodes contending for the same channel $i$ except for the successful competitor.

### 3.3 Utility function for hybrid channel access

Based on (6), we formulate a cost-type utility function to quantify the average delay experienced by a vehicular node in the group of $n_i$ nodes, which can indicate the congestion degree of their accessing channel $i$

$$\phi_i(n_i) = \frac{C}{r_i(n_i)}, \quad (7)$$

where $C > 0$ is a positive parameter that can be specified by the channel in order to quantify the magnitude of the channel congestion. For the sake of example here, we can simply set $C = 1$. Recalling that $r_i(n_i)$ represents the average number of packets expected to be successfully transmitted per time slot, $\phi_i(n_i)$ can denote the average transmission delay per packet. Based on (7) we can quantify the communication performance of vehicular nodes: if many vehicular nodes contend for the same channel $i$, i.e., $n_i$ is large, $i$ becomes congested, which can incur a high transmission delay, i.e., a large cost utility $\phi_i(n_i)$.

## 4 Evolutionary Game Formulation

### 4.1 Game Model

As shown in much existing literature such as [18], [28], evolutionary game theory can provide a refined Nash equilibrium with guaranteed evolutionary stability, i.e., evolutionarily stable strategy, and explicitly characterize the strategy-learning and adaptation of individuals, when compared to other additional non-cooperative game theories. It presents a powerful framework for modeling the dynamics of evolving strategies in a population, which has been widely applied in not only biology but also other engineering fields. Here, we consider the vehicular nodes are driven by their own interests, i.e., with the goal of minimizing their experienced cost utility, to independently make channel access decisions. Therefore, to take into account the bounded rationality of the vehicular nodes and capture the dynamics of interactions among them, we also resort to evolutionary game theory. To be specific, we model the access strategy adaptation of self-interest-driven vehicular nodes in a cognitive vehicular network as an evolutionary game, which formally consists of the following components:

- **Players**: In the evolutionary game, a player is a vehicular node that aims to minimize its experienced cost utility by adapting its decision on accessing a shared channel. Hence, the set of the players is $\mathcal{N}$.
- **Strategy**: A strategy of any player is the decision to access a shared channel in a competitive manner. At this point, the strategy set of each player can be $\mathcal{M}$.
- **Population**: A group of players that are sharing the same channel constitute a population. We also denote by $i \in \mathcal{M}$ the index of a population associated with the strategy $i$.
- **Population share**: The population share is referred to the proportion of the players in the population. Specifically, the population share is defined as $s_i = n_i/N$. Noting $N = \sum_{i=1}^{M} n_i$ and $n_i \in \mathbb{Z}_0^+$, we equivalently have $1 = \sum_{i=1}^{M} s_i$ and $\forall s_i \in [0, 1]$.
- **Population state**: We represent the state of the populations as the collection of all the population shares, i.e., $\boldsymbol{s} = [s_1, s_2, \ldots, s_M]^{\mathrm{T}}$.
- **Expected payoff**: The expected payoff of a player choosing the shared channel $i$ can be the cost utility associated with this channel. Given a population state $\boldsymbol{s} \in [0, 1]^M$, with equation (7) and the notation of the population share $s_i$ above we rearrange the cost-type utility as follows

$$f_i(s_i, \boldsymbol{s}) = \phi_i(Ns_i). \quad (8)$$

We remark that $f_i(s_i, \boldsymbol{s})$ is (i) a monotonously non-decreasing mapping with respect to the population share $s_i$, i.e., $f_i : [0, 1] \to \mathbb{R}_{\ge 0}$, which is also (ii) a twice continuously differentiable and (iii) convex function of $s_i$. The proof of these basic properties of $f_i(s_i, \boldsymbol{s})$ is given in the on-line supplementary

material. It can be seen that the expected payoff is also affected by the other players' strategies besides $s_i$ since $1 = \sum_{i=1}^{M} s_i$.

Subsequently, for a better understanding, we present the definition of a Nash equilibrium in the game formulation.

***Definition** 1: A population share profile, i.e., the population state $\boldsymbol{s}^{\mathrm{NE}}$, is a Nash equilibrium in the evolutionary game formulation if for all the nonzero population shares $s_i > 0$, $i \in \mathcal{M}$, the corresponding cost-type payoff satisfies $f_i(s_i, \boldsymbol{s}^{\mathrm{NE}}) \leq f_{i'}(s_{i'}, \boldsymbol{s}^{\mathrm{NE}})$ for $\forall i' \in \mathcal{M}$.*

From Definition 1, we can see in the Nash equilibrium state:

- **Remark 1**: Any utilized shared channel $i$ with a nonzero population share $s_i > 0$ must has an identical cost utility. Thus, the expected payoff of a player accessing the channel is equal to each other's, i.e., equal to the average level.
- **Remark 2**: The rest channels with a zero population share $s_{i'} = 0$ must provide a larger or equal cost utility.

**Remark 1** and **Remark 2** are due to the fact that none of the players would like to unilaterally change its channel access when they arrive at the Nash equilibrium.

### 4.2 Analysis of Nash equilibrium

A Nash equilibrium is a quite important solution to a noncooperative game. For the game formulated above, we have the following theorem to describe its basic properties related to the Nash equilibrium.

**Theorem 1.** *$\boldsymbol{s}^{\mathrm{NE}} = \left[s_1^{\mathrm{NE}}, s_2^{\mathrm{NE}}, \ldots, s_M^{\mathrm{NE}}\right]^{\mathrm{T}}$ is a Nash equilibrium if and only if it is characterized as the global minimizer of the following model*

$$\boldsymbol{s}^{\mathrm{NE}} \in \underset{\boldsymbol{s}\in[0,1]^M}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} \int_0^{s_i} f_i(x, \boldsymbol{s}) dx, 1 = \sum_{i=1}^{M} s_i \right\} \tag{9}$$

The proof of Theorem 1 is available in the online supplemental material.

Generally, the Nash equilibrium does not necessarily optimize the global performance of the overall game-theoretic system, as stated in existing literature [8], [9], [10], which is widely known as the inefficiency of the Nash equilibrium. To investigate the quality of the Nash equilibrium existing in the proposed game, we introduce an aggregate cost utility, called *the social cost*, as the global performance metric

$$G(\boldsymbol{s}) = \sum_{i=1}^{M} N s_i f_i(s_i, \boldsymbol{s}) = \sum_{i=1}^{M} \frac{C N s_i}{\alpha_i (1-\alpha_i)^{N s_i - 1} \overline{r}_i}, \tag{10}$$

where $\overline{r}_i = \sum_{l=1}^{L} c_{l,i} \mathrm{Prob}_1 \{c_{l,i}, i\} + \sum_{l'=1}^{L} c'_{l'} \mathrm{Prob}_2 \{c'_{l'}\}$. $G(\boldsymbol{s})$ is the total sum of all the players' cost, so it is also a convex function with respect to each $s_i$. Subsequently, a feasible optimal strategy profile, denoted by $\boldsymbol{s}^*$, can be defined as the socially optimal solution that minimizes the social cost over all the feasible strategies, i.e.,

$$\boldsymbol{s}^* \in \underset{\boldsymbol{s}\in[0,1]^M}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} N s_i f_i(s_i, \boldsymbol{s}), 1 = \sum_{i=1}^{M} s_i \right\} \tag{11}$$

Logically, a measure of the quality of the Nash equilibrium solution $\boldsymbol{s}^{\mathrm{NE}}$ can be formulated as the ratio between the social cost of the equilibrium solution $\boldsymbol{s}^{\mathrm{NE}}$ and that of the centrally designed optimal solution $\boldsymbol{s}^*$, $\theta = G\left(\boldsymbol{s}^{\mathrm{NE}}\right) / G\left(\boldsymbol{s}^*\right)$, which is known as *the price of anarchy* in economics and game theory [38]. Based on the definition, we derive some results on $\theta$.

**Lemma 1.** *Suppose that $\boldsymbol{s}^{\mathrm{NE}}$ is the Nash equilibrium of the game formulated in (9). For any feasible solution of (9), the following inequality always holds*

$$\sum_{i=1}^{M} s_i^{\mathrm{NE}} f_i\left(s_i^{\mathrm{NE}}, \boldsymbol{s}^{\mathrm{NE}}\right) \leq \sum_{i=1}^{M} \widetilde{s}_i f_i\left(s_i^{\mathrm{NE}}, \boldsymbol{s}^{\mathrm{NE}}\right). \tag{12}$$

**Lemma 2.** *Suppose that any two constants $\lambda$ and $\mu$ satisfy $\mu \in (0,1)$ and $\lambda\mu \geq \frac{1}{4}\max\left\{(1-\alpha_i)^{-N}, i = 1, 2, \ldots, M\right\}$. The following inequality always holds for any strategy $i$ and two population shares $s_i, s'_i \in [0,1]$*

$$s'_i f_i(s_i, \boldsymbol{s}) \leq \lambda s'_i f_i(s'_i, \boldsymbol{s}') + \mu s_i f_i(s_i, \boldsymbol{s}), \tag{13}$$

*where $\boldsymbol{s}'$ is the population state associated with $s'_i$.*

The proofs of Lemmas 1 and 2 are given in the online supplementary material. Now, with Lemmas 1 and 2, we can derive the theoretical bounds of the price of anarchy.

**Theorem 2.** *Suppose that any two constants $\lambda$ and $\mu$ satisfy $\mu \in (0,1)$ and $\lambda\mu \geq \frac{1}{4}\max\left\{(1-\alpha_i)^{-N}, i = 1, 2, \ldots, M\right\}$. If the feasible $\boldsymbol{s}^{\mathrm{NE}}$ is the Nash equilibrium of the game formulation (9) and $\boldsymbol{s}^*$ is an optimal solution with respect to the global optimization problem (11), then the price of anarchy in the game (9) is bounded as $1 \leq \theta \leq \frac{\lambda}{1-\mu}$.*

The proof of Theorem 2 is given in the online supplementary material. From Theorem 2, it is seen that the price of anarchy in the Nash equilibrium is related to the competition among vehicular nodes. The larger $\alpha_i$ is, the fiercer the competition becomes. Increasing the competition, i.e., $\alpha_i$, will increase the parameter $\lambda$, which leads to a higher upper bound of the price of anarchy.

## 5 Evolutionary Dynamics of Hybrid Channel Access with Dynamic Congestion Pricing

### 5.1 Dynamic Congestion Pricing

Let $q_i(s_i, \boldsymbol{s}) \in \mathbb{R}_{\geq 0}$ be the toll a vehicular node has to pay for accessing the shared channel $i$ given that the population share associated with $i$ is $s_i$ and the current population state is $\boldsymbol{s}$. Specifically, we refer to the formulation of the *marginal cost taxes* first proposed by Pigou [11] to devise $q_i(s_i, \boldsymbol{s})$, which can be expressed as $q_i(s_i, \boldsymbol{s}) = N s_i \frac{df_i(s_i, \boldsymbol{s})}{dN s_i} = s_i \frac{df_i(s_i, \boldsymbol{s})}{ds_i}$:

$$q_i(s_i, \boldsymbol{s}) = \frac{-C N s_i \ln(1-\alpha_i)}{\alpha_i (1-\alpha_i)^{N s_i - 1} \overline{r}_i}. \tag{14}$$

The marginal cost pricing is believed to be a powerful approach to reducing the inefficiency of the Nash equilibrium [38], the underlying idea of which is to charge each player for an additional cost incurred by its presence in the spectrum competition. Now, with the marginal cost pricing $q_i(s_i, \boldsymbol{s})$, any vehicular node in the $i$-th population must suffer the total cost of $f_i(s_i, \boldsymbol{s}) + q_i(s_i, \boldsymbol{s})$. At this point,

we can modify the cost utility experienced by a player in the formulated game to $\pi_i(s_i, \boldsymbol{s}) = f_i(s_i, \boldsymbol{s}) + q_i(s_i, \boldsymbol{s})$ for $\forall i \in \mathcal{M}$. Obviously, $q_i(s_i, \boldsymbol{s})$ is also twice continuously differentiable with respect to $s_i$. We can further derive $f_i'''(s_i, \boldsymbol{s}) = \frac{-CN^3 \ln^3(1-\alpha_i)}{\alpha_i(1-\alpha_i)^{Ns_i-1}\overline{r}_i} > 0$ under $\alpha_i \in (0,1)$. Noting $q_i''(s_i, \boldsymbol{s}) = 2f_i''(s_i, \boldsymbol{s}) + s_i f_i'''(s_i, \boldsymbol{s}) > 0$, we see that $q_i(s_i, \boldsymbol{s})$ is strictly convex with respect to $s_i$. This also indicates the strict convexity of the total cost utility of each individual player, $\pi_i(s_i, \boldsymbol{s})$. Thus, similar to (9), the Nash equilibrium of the evolutionary game with the marginal pricing, $\boldsymbol{z}^{\mathrm{NE}}$, can be specified as the minimizer of the following convex optimization model

$$\boldsymbol{z}^{\mathrm{NE}} \in \underset{\boldsymbol{s}\in[0,1]^M}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} \int_0^{s_i} \pi_i(x, \boldsymbol{s}) dx, 1 = \sum_{i=1}^{M} s_i \right\} \tag{15}$$

Furthermore, we establish the connection between $\boldsymbol{z}^{\mathrm{NE}}$ and the socially optimal solution of (11), $\boldsymbol{s}^*$.

**Theorem 3.** *A strategy profile, $\boldsymbol{z}^{\mathrm{NE}}$, is the Nash equilibrium of the evolutionary game with the dynamic congestion pricing as represented by (15) if and only if it is the socially optimal solution of the model (11). Meanwhile, given that $\boldsymbol{s}^*$ is socially optimal with respect to (11), $\boldsymbol{s}^*$ must be the Nash equilibrium solution of (15) with imposing the channel congestion prices $q_i(s_i^*, \boldsymbol{s}^*)$ for $\forall i \in \mathcal{M}$.*

The proof of Theorem 3 is given in the online supplementary material. Furthermore, to analyze the evolutionary stability of the Nash equilibrium strategy, we refer to the concept of evolutionary stable strategy (ESS) [6], [7]. Suppose that a small group of the players, the proportion of which is denoted by $\chi \in (0,1)$, change their currently adopted strategies $\boldsymbol{s}$ to other alternatives. The profile of these mutants' new strategies can be denoted by $\boldsymbol{s}_\chi \in [0,1]^M$ and $1 = \sum_{i=1}^{M} s_{i,\chi}$ where $s_{i,\chi}$ is the ratio of the mutants that switch to access the shared channel $i$ to the total mutants. In this sense, the new population state after the mutation can be expressed as $(1-\chi)\boldsymbol{s} + \chi\boldsymbol{s}_\chi$. Let $\overline{\pi}(\boldsymbol{s})$ be the the average cost utility under $\boldsymbol{s}$. The ESS of the game model (15) can be mathematically defined as follows.

***Definition** 2 [7]: $\boldsymbol{s}^{\mathrm{ESS}}$ is an ESS of (15) if for any $\boldsymbol{s}_\chi \neq \boldsymbol{s}^{\mathrm{ESS}}$, there is a certain $\widetilde{\chi} \in (0,1)$ such that $\overline{\pi}\left(\boldsymbol{s}^{\mathrm{ESS}}\right) < \overline{\pi}\left((1-\chi)\boldsymbol{s}^{\mathrm{ESS}} + \chi\boldsymbol{s}_\chi\right)$ is always held for $\forall \chi \in (0, \widetilde{\chi})$.*

The ESS represents an equilibrium refinement of the Nash equilibrium strategy that is able to prevent some mutant strategies from invading if the mutant strategies are rare at the initialization of the population evolution. Now, we can establish the relationship between the Nash equilibrium of the game (15) and the ESS as follows.

**Theorem 4.** *The Nash equilibrium strategy $\boldsymbol{z}^{\mathrm{NE}}$ obtained from (15) must also be the evolutionary stable strategy $\boldsymbol{s}^{\mathrm{ESS}}$.*

The proof of Theorem 4 is available in the online supplementary material.

Besides, by applying Theorems 3 and 4, we can derive that the socially optimal solution to (11) is also an ESS.

### 5.2 Evolutionary Dynamics of Strategy Adaptation

Let $t \in \mathbb{R}_{\geq 0}$ be the time index where $t = 0$ indicates the initialization of the evolutionary game. Without ambiguity, we can slightly modify the notations $s_i$ and $\boldsymbol{s}$ to $s_i(t)$ and $\boldsymbol{s}(t)$, which express the population share and the population state at time $t$, respectively. For simplicity, we introduce the following $(M-1)$-dimensional standard simplex $\Delta^{M-1}$:

$$\Delta^{M-1} = \left\{ [s_1, s_2, \ldots, s_M]^{\mathrm{T}} \,\middle|\, \forall s_i \in [0,1] \text{ and } \sum_{i=1}^{M} s_i = 1 \right\} \tag{16}$$

to represent the constraints of the model (15). Denoting by $\overline{\pi}(\boldsymbol{s}(t))$ the average cost of the overall population at time $t$, $\overline{\pi}(\boldsymbol{s}(t)) = \sum_{i=1}^{M} s_i(t)\pi_i(s_i(t), \boldsymbol{s}(t))$, we can use the replicator dynamics [7] to model the evolutionary game with the dynamic channel pricing corresponding to (15), which is defined on the simplex $\Delta^{M-1}$ as follows:

$$\frac{ds_i(t)}{dt} = \sigma s_i(t) \left[\overline{\pi}(\boldsymbol{s}(t)) - \pi_i(s_i(t), \boldsymbol{s}(t))\right] \tag{17}$$

with the initial population state $\boldsymbol{s}(0) \in \Delta^{M-1}$, for $\forall i \in \mathcal{M}$, where $\sigma$ is the learning rate that controls the frequency of the strategy adaptation for channel accessing. For $\sigma$, we assume

$$0 < \sigma < \frac{1}{\max_{\boldsymbol{s}\in\Delta^{M-1}}\{\pi_i(s_i, \boldsymbol{s})\}}. \tag{18}$$

In fact, such a positive $\sigma$ always exists because of the continuity of the cost utility function $\pi_i(s_i, \boldsymbol{s})$ with respect to $s_i$ and the boundedness of $s_i$ for all $i \in \mathcal{M}$.

From (17), we can see that a positive population share $s_i(t) > 0$ will be decreased with a negative growth rate $s_i'(t) < 0$ if the corresponding cost is larger than the average cost, i.e., $\pi_i(s_i(t), \boldsymbol{s}(t)) > \overline{\pi}(\boldsymbol{s}(t))$, while it will increase if $s_i'(t) > 0$. When $s_i(t) = 0$, it will stop evolving since its growth rate is always zero, i.e., $s_i'(t) = 0$. The evolutionary equilibria (EE) of the dynamical system (17) may be considered as the solution of the system. The EE are usually defined as the fixed points of the dynamical system, i.e., the points satisfying $ds_i(t)/dt = 0 \ \forall s_i(t) \in \boldsymbol{s}(t)$. At the fixed points, none of the population shares change any longer and the fairness of spectrum allocation among the vehicular nodes can also be achieved. We let the set of the EE of (17) be $\mathcal{U}_{\mathrm{EE}}$, $\mathcal{U}_{\mathrm{EE}} = \left\{\boldsymbol{s}(t) \in \Delta^{M-1} \,|ds_i(t)/dt = 0 \ \forall i \in \mathcal{M}\right\}$. Unfortunately, not all of the EE are stable. In other words, the unstable points, if existing in $\mathcal{U}_{\mathrm{EE}}$, are not the ESS. Thus, they must not be the Nash equilibrium of the model (15), i.e., not the socially optimal solution to (11), according to Theorems 3 and 4. Formally, we define a set $\widetilde{\mathcal{U}}_{\mathrm{EE}}$ by $\widetilde{\mathcal{U}}_{\mathrm{EE}} = \left\{\boldsymbol{s}(t) \in \mathcal{U}_{\mathrm{EE}} \,|\exists i \ s_i(t) = 0 \text{ and } \pi_i\left(s_i(t), \boldsymbol{s}(t)\right) < \overline{\pi}\left(\boldsymbol{s}(t)\right)\right\}$. It is obvious that all the points in $\widetilde{\mathcal{U}}_{\mathrm{EE}}$ are neither stable due to the fact that any small perturbation will make the system deviate from the equilibrium state, nor the Nash equilibrium according to Definition 1. Denote the set of the socially optimal solutions to (11) by $\mathcal{U}_{\mathrm{SO}}$, i.e., $\mathcal{U}_{\mathrm{SO}} = \operatorname{argmin}_{\boldsymbol{s}\in\Delta^{M-1}}\{G(\boldsymbol{s})\}$. We connect the EE of the replicator dynamics (17) with the socially optimal solution to (11) by the following theorem:

**Theorem 5.** *$\mathcal{U}_{\mathrm{SO}} = \mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$ always holds for (17).*

The proof of Theorem 5 is available in the online supplementary material. By following Theorem 5, we further have the following results:

- **Remark 3**: $\left|\mathcal{U}_{\mathrm{EE}}\backslash\widetilde{\mathcal{U}}_{\mathrm{EE}}\right| = 1$, because the strict convexity of (11) guarantees the existence of a unique global optimal solution in its feasible region, i.e., $|\mathcal{U}_{\mathrm{SO}}| = 1$.
- **Remark 4**: $\mathcal{U}_{\mathrm{SO}}$ is an invariant manifold of $\Delta^{M-1}$, as $\mathcal{U}_{\mathrm{EE}}\backslash\widetilde{\mathcal{U}}_{\mathrm{EE}}$ are invariant manifolds.
- **Remark 5**: $\mathcal{U}_{\mathrm{EE}}\backslash\widetilde{\mathcal{U}}_{\mathrm{EE}}$ is the set of the ESS of (17) according to Theorem 4.

Additionally, we show that the EE in $\mathcal{U}_{\mathrm{EE}}\backslash\widetilde{\mathcal{U}}_{\mathrm{EE}}$ is asymptotically stable as follows.

**Theorem 6.** *The socially optimal solution $\boldsymbol{s}^* \in \mathcal{U}_{\mathrm{SO}} = \mathcal{U}_{\mathrm{EE}}\backslash\widetilde{\mathcal{U}}_{\mathrm{EE}}$ of the model (11) is asymptotically stable and there exists a sufficiently small $\bar{\epsilon}$ such that for all $0 < \epsilon < \bar{\epsilon}$, from any given initial point $\boldsymbol{s}(0) \in \Delta^{M-1}$ which satisfies $\boldsymbol{s}(0) \neq \boldsymbol{s}^*$ and $\|\boldsymbol{s}(0) - \boldsymbol{s}^*\| < \epsilon$ where $\|\cdot\|$ denotes the Euclidean distance operator, the replicator dynamics (17) can converge to $\boldsymbol{s}^*$.*

The proof of Theorem 6 is available in the online supplementary material.

### 5.3 Discretized Evolutionary Dynamics with Delayed Pricing

To practically implement the strategy adaptation with the replicator dynamics formulated above, we further consider to discretize (17) by introducing $k \in \mathbb{Z}_0^+$ as the index of each channel access decision time slot (i.e., each strategy adaptation epoch). Thus, it is assumed that the game of competitively accessing hybrid channels is repeatedly played by vehicular nodes at each epoch $k$ on an infinite time horizon (See Fig. 1). At each $k$, vehicular nodes can access their preferred shared channels, which incurs a population distribution of the entire nodes, $\boldsymbol{s}(k) = [s_1(k), \ldots, s_M(k)]^{\mathrm{T}}$. Then, these vehicular nodes observe the cost-type payoff relevant to their decisions, including the actual channel performance cost $\{f_i(s_i(k), \boldsymbol{s}(k)), \forall i \in \mathcal{M}\}$ and the additionally imposed channel congestion prices $\{q_i(s_i(k), \boldsymbol{s}(k)), \forall i \in \mathcal{M}\}$. Using the cost information, they can adapt their access decisions at the subsequent epoch $k+1$. Then, the discretized realization of (17) can be expressed as

$$s_i(k+1) = s_i(k) + \sigma(k)s_i(k)\left[\overline{\pi}(\boldsymbol{s}(k)) - \pi_i(s_i(k), \boldsymbol{s}(k))\right]. \tag{19}$$

In (19), two types of the learning rate are investigated, one of which corresponds to the time-dependent learning rate and the other the constant rate. For the first case, we denote the set of a series of the learning rates at each epoch by $\Theta_1 = \left\{\sigma(k) = \sigma\theta(k) \in \mathbb{R}_{>0} \,\middle|\, \sum_{k=0}^{\infty}\theta(k) = \infty, \sum_{k=0}^{\infty}\theta^2(k) < \infty\right\}$. In this case, the learning rate $\sigma(k)$ is vanishing over iterations $k$. In contrast, for the constant rate, we denote $\Theta_2 = \{\sigma(k) = \sigma\theta(k) \in \mathbb{R}_{>0} \,|\, \theta(k) = \theta\ \forall k\}$ where $\theta$ is a small positive real number.

On the other hand, let $W \in \mathbb{Z}^+$ be the size of the time period for updating the channel congestion prices and $W_k = k - W \times \left\lfloor \frac{k}{W} \right\rfloor$ be the difference between the current epoch of strategy adaptation and that of congestion pricing, where $\lfloor\cdot\rfloor$ is the floor function. We delay the marginal cost pricing formula (14) by $W_k$ and modify the total cost utility $\pi_i(s_i(k), \boldsymbol{s}(k)) = f_i(s_i(k), \boldsymbol{s}(k)) + q_i\left(s_i(k), \boldsymbol{s}(k)\right)$ to

$$\begin{aligned}\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k)) =& f_i(s_i(k), \boldsymbol{s}(k)) \\ &+ q_i\left(s_i\left(k - W_k\right), \boldsymbol{s}\left(k - W_k\right)\right).\end{aligned} \tag{20}$$

Correspondingly, the average payoff $\overline{\pi}(\boldsymbol{s}(k))$ is changed to $\overline{\pi}^{W_k}(\boldsymbol{s}(k)) = \sum_{i=1}^{M} s_i(k)\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$.

Based on (17), the discretized iterative equation of the replicator dynamics with delayed dynamic channel congestion pricing for $k \in \mathbb{Z}_0^+$ can be represented by

$$s_i(k+1) = s_i(k) + \sigma(k)s_i(k)\left[\overline{\pi}^{W_k}(\boldsymbol{s}(k)) - \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))\right] \tag{21}$$

with the initialization $\boldsymbol{s}(0) \in (0,1)^M$ and the learning rate $\sigma(k) \in \Theta_1$ or $\sigma(k) \in \Theta_2$.

### 5.4 Convergence Analysis

With the pricing delay $W_k$, the discretized evolutionary dynamics (21) may be different from that of the original system (19). To shed light on the effect of $W_k$ on the discretized evolutionary dynamics, we first provide some basic properties and then establish the connection between the discretized evolutionary dynamics with the delayed pricing and the original one without a delay in pricing.

**Lemma 3.** *There always exists a positive constant $Q \in \mathcal{R}_{>0}$ such that $\max\left\{\left|\partial q_i(s_i, \boldsymbol{s})/\partial s_l\right|,\ \forall i, l \in \mathcal{M}\right\} \leq Q$ holds.*

**Lemma 4.** *Given a $W \in \mathbb{Z}^+$, $W_k \leq W$ always holds.*

**Lemma 5.** *Let $\sigma s_i(k)\left[\overline{\pi}^{W_k}(\boldsymbol{s}(k)) - \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))\right] = \Delta s_i(k)$ for all $i \in \mathcal{M}$ and then $|\Delta s_i(k)| \leq s_i(k)$ always holds.*

The proofs of Lemmas 3-5 are available in the online supplementary material. Based on these lemmas, we show that the difference between the pricing mechanisms with and without a delay is bounded and the bound is in the same order of the adopted learning rate.

**Theorem 7.** *There always exists a positive constant $\vartheta \in \mathbb{R}_{>0}$ such that for all $i \in \mathcal{M}$ the following holds for either $\sigma(k) \in \Theta_1$ or $\sigma(k) \in \Theta_2$*

$$\begin{aligned}&\left|q_i\left(s_i(k), \boldsymbol{s}(k)\right) - q_i\left(s_i(k - W_k), \boldsymbol{s}(k - W_k)\right)\right| \\ &\leq \vartheta \sum_{\tau=0}^{W-1}\theta(k - \tau - 1).\end{aligned} \tag{22}$$

*Proof:* For simplicity, we still use the notation of $\Delta s_i(k)$ in Lemma 5 and also define a column vector of $\Delta s_i(k)$ by $\Delta\boldsymbol{s}(k) = [\Delta s_1(k), \ldots, \Delta s_M(k)]^{\mathrm{T}}$. Based on (21) and noticing the definition of the learning rate $\sigma(k) = \sigma\theta(k)$, we can directly yield $q_i\left(s_i(k+1), \boldsymbol{s}(k+1)\right) = q_i\left(s_i(k) + \theta(k)\Delta s_i(k), \boldsymbol{s}(k) + \theta(k)\Delta\boldsymbol{s}(k)\right)$. Regarding the small parameter $\theta(k)$ as a scalar variable and the marginal cost utility as the function of $\theta(k)$, i.e., letting $\varphi_i(\xi) = q_i\left(s_i(k) + \xi\Delta s_i(k), \boldsymbol{s}(k) + \xi\Delta\boldsymbol{s}(k)\right)$, we can obtain the linear expansion of the marginal cost $\varphi_i(\xi)$ at the zero point $\xi = 0$ by using Taylor's theorem as follows

$$\varphi_i(\xi) = \varphi_i(0) + \varphi_i'(\tilde{\xi})\xi, \tag{23}$$

where $\tilde{\xi}$ is some real number between 0 and $\xi$. In addition, the first-order derivative of $\varphi_i(\xi)$ at $\tilde{\xi}$ can be expressed as $\varphi_i'(\tilde{\xi}) = \sum_{l=1}^{M}\left(\partial q_i(s_i^{\tilde{\xi}}, \boldsymbol{s}^{\tilde{\xi}})/\partial s_l\right)\Delta s_l(k)$, where we denote

$s_i^{\tilde{\xi}} = s_i(k) + \tilde{\xi}\Delta s_i(k)$ and $\boldsymbol{s}^{\tilde{\xi}} = \boldsymbol{s}(k) + \tilde{\xi}\Delta \boldsymbol{s}(k)$. Thus, according to Lemmas 3 and 5, we can see

$$\left|\varphi_i'(\tilde{\xi})\right| \leq \sum_{l=1}^{M}\left|\frac{\partial q_i(s_i^{\tilde{\xi}}, \boldsymbol{s}^{\tilde{\xi}})}{\partial s_l}\right| \times |\Delta s_l(k)| \leq \sum_{l=1}^{M} Q s_l(k) = Q. \tag{24}$$

(24) indicates that there must exist some real numbers $\eta_i(k)$, satisfying $|\eta_i(k)| \leq Q$, such that $q_i\left(s_i(k+1), \boldsymbol{s}(k+1)\right) = q_i\left(s_i(k), \boldsymbol{s}(k)\right) + \eta_i(k)\theta(k)$. That is, $|q_i\left(s_i(k+1), \boldsymbol{s}(k+1)\right) - q_i\left(s_i(k), \boldsymbol{s}(k)\right)| \leq Q\theta(k)$. Using this result and Lemma 4, we can reach

$$\begin{aligned}
&|q_i\left(s_i(k), \boldsymbol{s}(k)\right) - q_i\left(s_i(k-W_k), \boldsymbol{s}(k-W_k)\right)| \\
&= \left|\sum_{\tau=0}^{W_k-1}\begin{bmatrix} q_i\left(s_i(k-\tau), \boldsymbol{s}(k-\tau)\right) \\ -\, q_i\left(s_i(k-\tau-1), \boldsymbol{s}(k-\tau-1)\right)\end{bmatrix}\right| \\
&\leq \sum_{\tau=0}^{W-1}\left|\begin{bmatrix} q_i\left(s_i(k-\tau), \boldsymbol{s}(k-\tau)\right) \\ -\, q_i\left(s_i(k-\tau-1), \boldsymbol{s}(k-\tau-1)\right)\end{bmatrix}\right| \\
&\leq Q\sum_{\tau=0}^{W-1}\theta(k-\tau-1).
\end{aligned} \tag{25}$$

By selecting $\vartheta$ as $Q$, the proof is completed. ■

Following Theorem 7, we can show the convergence of the discretized evolutionary dynamics with delayed pricing.

**Theorem 8.** *Suppose that the series* $\left\{\boldsymbol{s}(k)\ \forall k \in \mathbb{Z}^+_{\geq W_k}\right\}$ *is generated from the discretized evolutionary dynamics with the delayed pricing (21). Then, we can have*

$$\begin{aligned}
s_i(k+1) =& s_i(k) \\
&+\sigma(k)s_i(k)\left[\overline{\pi}(\boldsymbol{s}(k)) - \pi_i(s_i(k), \boldsymbol{s}(k))\right] + \mathcal{O}\left(\theta^2(k)\right)
\end{aligned} \tag{26}$$

*for either* $\sigma(k) \in \Theta_1$ *or* $\sigma(k) \in \Theta_2$.

*Proof:* Using the notation of $\Delta s_i(k)$ defined in Lemma 5 and defining $\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k)) - \pi_i(s_i(k), \boldsymbol{s}(k)) = q_i(s_i(k-W_k), \boldsymbol{s}(k-W_k)) - q_i(s_i(k), \boldsymbol{s}(k)) = \Delta q_i(k|W_k)$, we can rearrange (21) as follows

$$\begin{aligned}
s_i(k+1) =& s_i(k) \\
&+ \sigma\theta(k)s_i(k)\left[\overline{\pi}^{W_k}(\boldsymbol{s}(k)) - \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))\right] \\
&+ \sigma\theta(k)s_i(k)\left[\overline{\pi}(\boldsymbol{s}(k)) - \pi_i(s_i(k), \boldsymbol{s}(k))\right] \\
&- \sigma\theta(k)s_i(k)\left[\overline{\pi}(\boldsymbol{s}(k)) - \pi_i(s_i(k), \boldsymbol{s}(k))\right] \\
=& s_i(k) + \sigma\theta(k)s_i(k)\left[\overline{\pi}(\boldsymbol{s}(k)) - \pi_i(s_i(k), \boldsymbol{s}(k))\right] \\
&+ \sigma\theta(k)s_i(k)\left\{\sum_{l=1}^{M} s_l(k)\Delta q_l(k|W_k) - \Delta q_i(k|W_k)\right\}.
\end{aligned} \tag{27}$$

Next, based on Theorem 7 and noticing $\sum_{l=1}^{M} s_l(k) = 1$ and $\forall s_i(k) \in [0,1]$, we can see

$$\begin{aligned}
&\left|\sigma\theta(k)s_i(k)\left\{\sum_{l=1}^{M} s_l(k)\Delta q_l(k|W_k) - \Delta q_i(k|W_k)\right\}\right| \\
&\leq \sigma\theta(k)s_i(k)\left\{\sum_{l=1}^{M} s_l(k)\left|\Delta q_l(k|W_k)\right| + \left|\Delta q_i(k|W_k)\right|\right\} \\
&\leq 2\sigma\vartheta\sum_{\tau=0}^{W-1}\theta(k)\theta(k-\tau-1),
\end{aligned} \tag{28}$$

which indicates that $\sigma\theta(k)s_i(k)\{\sum_{l=1}^{M} s_l(k)\Delta q_l(k|W_k) - \Delta q_i(k|W_k)\} = \mathcal{O}\left(\theta^2(k)\right)$. Thus, we complete the proof. ■

Theorem 8 reveals that the difference between the proposed evolutionary dynamics with the delayed pricing and that without a delay in the channel pricing is also bounded and the bound is in the same order of a square of the adopted learning rate, i.e., the order of $\theta^2(k)$. Moreover, recalling the definitions of $\Theta_1$ and $\Theta_2$, we can also see

- **Remark 6**: When $\sigma(k) \in \Theta_1$, the evolutionary dynamics with delayed pricing (21) can converge to the phase trajectory of (19) that is without any delay in the channel congestion prices as $k \to \infty$. This is because a vanishing learning rate can lead to $\mathcal{O}\left(\theta^2(k)\right) \to 0$ as $k \to \infty$.
- **Remark 7**: When $\sigma(k) \in \Theta_2$, $\theta(k)$ is specified as a constant $\theta$. In this case, we can see $|\sigma\theta(k)s_i(k)\{\sum_{l=1}^{M} s_l(k)\Delta q_l(k|W_k) - \Delta q_i(k|W_k)\}| \leq 2W\sigma\vartheta\theta^2$. This means that the phase trajectory of (21) is bounded near that of (19), and the bound of the difference between these two phase trajectories is in the order of $\theta^2$.

Furthermore, to investigate the steady performance of the discretized evolutionary dynamics (21) with $\sigma(k) \in \Theta_1$ and $\sigma(k) \in \Theta_2$, respectively, we illustrate anther mathematical property of $\pi_i(s_i, \boldsymbol{s})$ in the following lemma.

**Lemma 6.** *$\pi_i(s_i, \boldsymbol{s})$ is Lipschitz continuous with respect to $s_i$ (or $\boldsymbol{s}$) over $\Delta^{M-1}$ $\forall i \in \mathcal{M}$.*

Now, we can obtain the following theorems.

**Theorem 9.** *For $\sigma(k) \in \Theta_1$, the series $\{\boldsymbol{s}(k), \forall k \in \mathbb{Z}^+_{\geq W_k}\}$ generated from (21) can asymptotically converge to the socially optimal point $\boldsymbol{s}^*$ of the model (11) with an initial interior point $\boldsymbol{s}(0) \in (0,1)^M$. That is, $\lim_{k\to\infty}\|\boldsymbol{s}(k) - \boldsymbol{s}^*\| = 0$ from the interior region of $\Delta^{M-1}$.*

*Proof:* In fact, the Lipschitz continuity of the cost utility function $\pi_i(s_i, \boldsymbol{s})$ illustrated in Lemma 6 can ensure that the original ordinary differential equation (17) has a unique solution for any initial point $\boldsymbol{s}(0)$, i.e., depending continuously on $\boldsymbol{s}(0)$. As shown in Theorem 6, given an interior point of the simplex $\Delta^{M-1}$ as the initial point, (17) can converge to $\boldsymbol{s}^*$. Here, we follow the analysis logic in [39] to establish an appropriate continuous phase trajectory $\overline{\boldsymbol{s}}(t)$ $t \in \mathbb{R}_{\geq 0}$ from $\boldsymbol{s}(k)$ by using linear interpolation, and further show that its asymptotic convergence to the socially optimal solution $\boldsymbol{s}^*$ similar to the behavior of (17).

Denote a series of time instants by $\{t(0) = 0, t(k) = \sum_{\tau=0}^{k-1}\theta(\tau), k \leq 1\}$. As given $\sigma(k) \in \Theta_1$, we have $t(k) \to \infty$ and $\theta(k) = t(k+1) - t(k) \to 0$ as $k \to \infty$. We also let $\Delta t(k) = [t(k), t(k+1)]$ for all $k$. Thus, we can set the points $\overline{\boldsymbol{s}}(t(k)) = \boldsymbol{s}(k)$ for all $k$ and define $\overline{\boldsymbol{s}}(t)$ for $t \in \Delta t(k)$ with linear interpolation on the interval $\Delta t(k)$ as follows

$$\overline{\boldsymbol{s}}(t) = \boldsymbol{s}(k) + \left(\boldsymbol{s}(k+1) - \boldsymbol{s}(k)\right)\frac{t - t(k)}{t(k+1) - t(k)}, t \in \Delta t(k). \tag{29}$$

Given any positive real number $t_0 \in \mathbb{R}_{>0}$ and let $\boldsymbol{s}^{t_0}(t) = \left[s_1^{t_0}(t), \ldots, s_M^{t_0}(t)\right]^{\mathrm{T}}$ for $t \geq t_0$ be the unique solution to (17)

TABLE 1
Parameters for adaptive transmission over different channels

| Parameters | Values |
|---|---|
| The SNR thresholds $\Gamma_l$ | $\{0, 8, 11, 15, 25, 36, +\infty\}$ (dB) |
| Data rates for $\kappa = 1$ | $\{3, 4.5, 6, 12, 18, 27\}$ (Mbps) |
| Data rates for $\kappa = 2$ | $\{6, 9, 12, 24, 36, 54\}$ (Mbps) |
| Parameters $a_l, g_l$ | $\{1, 0.5, 0.4, 0.3, 0.2, 0.1\}$ |

starting at time instant $t_0$, i.e.,

$$\frac{ds_i^{t_0}(t)}{dt} = \sigma s_i^{t_0}(t)\left[\sum_{l=1}^{M} s_l^{t_0}(t)\pi_l\left(s_l^{t_0}(t), \boldsymbol{s}^{t_0}(t)\right) - s_i^{t_0}(t)\right] \tag{30}$$

with $\boldsymbol{s}^{t_0}(t_0) = \overline{\boldsymbol{s}}(t_0)$ and $t \geq t_0$.

Combining Theorem 8 and the fact $\mathcal{O}\left(\theta^2(k)\right) \to 0$ as $k \to \infty$ when $\sigma(k) \in \Theta_1$, and according to [39] (See Lemma 1 in Section 2 of [39]), we can see

$$\lim_{t_0 \to \infty} \sup_{t \in [t_0, t_0+T]} \left\|\overline{\boldsymbol{s}}(t) - \boldsymbol{s}^{t_0}(t)\right\| = 0 \tag{31}$$

for any given $T > 0$. This means that when $k \to \infty$, the discretized trajectory of (21) can asymptotically converge to that of (30). Notice that (30) and (17) follow the same update rule. In other words, the limiting behavior of (21) is similar to (17). Based on Theorem 6, we can conclude that (21), like (17), can asymptotically converge to $\boldsymbol{s}^* \in \mathcal{U}_{\text{SO}}$. ■

In contrast, with $\sigma(k) \in \Theta_2$, we have the following theorem.

**Theorem 10.** *For $\sigma(k) \in \Theta_2$, the series $\{\boldsymbol{s}(k), \forall k \in \mathbb{Z}^+_{\geq W_k}\}$ generated from (21) can asymptotically converge to a neighborhood of $\boldsymbol{s}^* \in \mathcal{U}_{\text{SO}}$ of the model (11) with an initial interior point $\boldsymbol{s}(0) \in (0,1)^M$. That is, $\lim_{k\to\infty} \|\boldsymbol{s}(k) - \boldsymbol{s}^*\| = \mathcal{O}(\theta)$ from the interior region of $\Delta^{M-1}$.*

*Proof:* Notice that a constant $\sigma(k) = \sigma\theta$ is adopted here for all $k$. This theorem can also be proven by using the same analysis logic as in Theorem 9 and by referring to [39] (See Theorem 3 in Section 9 of [39]). ■

Theorems 9 and 10 reveals a trade-off to specify a learning rate for the discretized evolutionary dynamics. A constant learning rate-based evolutionary dynamics may be better adaptive to a dynamic environment at the price of sacrificing the optimality of the steady state. Theoretically, a vanishing learning rate can guarantee the convergence of the discretized evolutionary dynamics to the socially optimal state, while it may be unsuitable for a dynamic situation in reality, since the convergence will slow down when the learning rate is approaching zero.

## 6 Performance Evaluation

### 6.1 Parameter Setting

We provide an implementation framework of the channel access optimization based on the proposed evolutionary dynamics with the delayed channel pricing in a practical scenario as is illustrated in Fig. 2. We consider a cognitive vehicular network in a service area with $M = 3$ radio channels ($\mathcal{M} = \{1, 2, 3\}$) commonly shared by a licensed user and with one exclusive-use channel. The fading parameter $m$ is set to $m = 1.0$ for the shared channels while $m = 1.5$

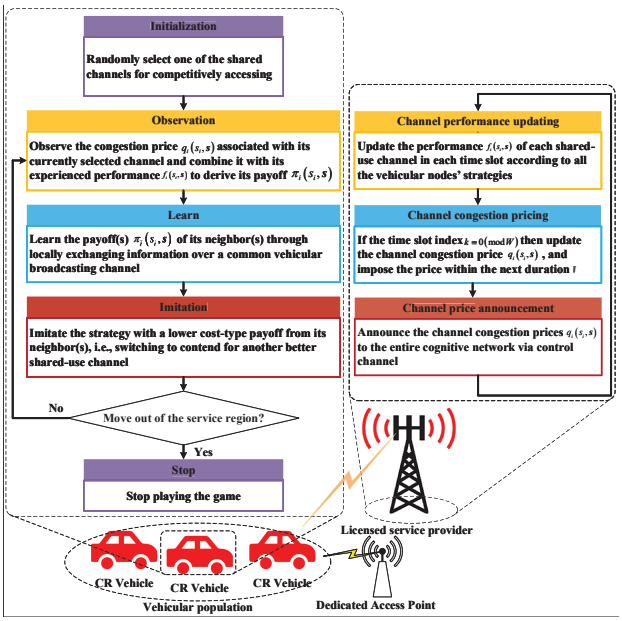



Fig. 2. A schematic diagram for implementation of the proposed evolutionary dynamics with the delayed dynamic pricing.

for the exclusive-use channel. The transmission probability of a vehicular node over each shared-use channel in a time slot is set as $\alpha_1 = 0.02$, $\alpha_1 = 0.03$, and $\alpha_1 = 0.05$. The channel qualities are given as $\overline{\omega}_1 \in \{11, 15, 19\}$ (dB) while $\overline{\omega}_2 = 12$ dB. Each packet size is set to 100 Bytes and the duration of each time slot is 20 ms. The average packet arrival rate is assumed to be $C = 1.0$ (packet per time slot). According to [37], the modulation of the DSRC-based exclusive-use channel can be based on different protocols such as BPSK, QPSK, 16-QAM, 64-QAM under the different coding rates $1/2, 2/3, 3/4$. [37] also shows that the parameter $b_l$ involved in equation (3) can be further approximated by $b_l = m/\overline{\omega}_\kappa + g_l$ where $g_l$ is the fitting parameter in mode $l$. Accordingly, in our experiments, we specify the transmission rates of channels under different modes as well as the SNR thresholds associated with each transmission mode in Table 1. It is worth pointing out that the parameter setting used here is for the sake of demonstration and our proposed model can also be applied in other scenarios.

### 6.2 Numerical Results

#### 6.2.1 *Evolutionary dynamics without channel pricing*

We numerically study the evolutionary dynamics without a pricing mechanism, for which we fix the population size $N$ at $N = 60$ and the population state is initialized as $\boldsymbol{s}(0) = [0.5, 0.2, 0.3]^{\text{T}}$. We select the dynamic learning rate as $\theta(k) = 100/(100 + k)$ for $\forall k$ and the constant learning rate as $\theta = 1$. The total epoch number is set to 5000. Fig. 3 shows the evolution of the population state $\boldsymbol{s}(k)$ under the two learning rates, while Fig. 4 shows the evolution of the cost-type payoff, $f_i(s_i, \boldsymbol{s})$, of the vehicular populations accessing different channels. From this two figures, we can observe that the vehicular populations can converge to the Nash equilibrium state where no vehicular node has the incentive to unilaterally adjust its access strategy as the payoff obtained from each utilized channel is identical. We can

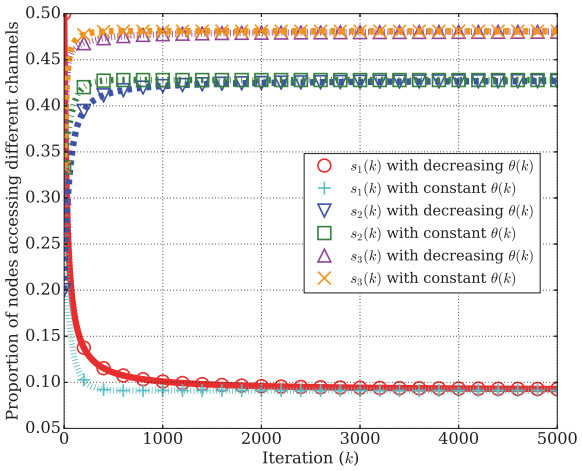



Fig. 3. Evolution of population shares without a channel congestion pricing.

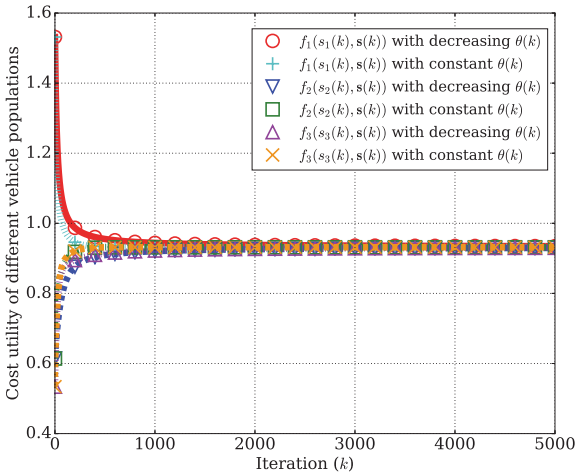



Fig. 4. Evolution of populations' payoffs without a channel congestion pricing.

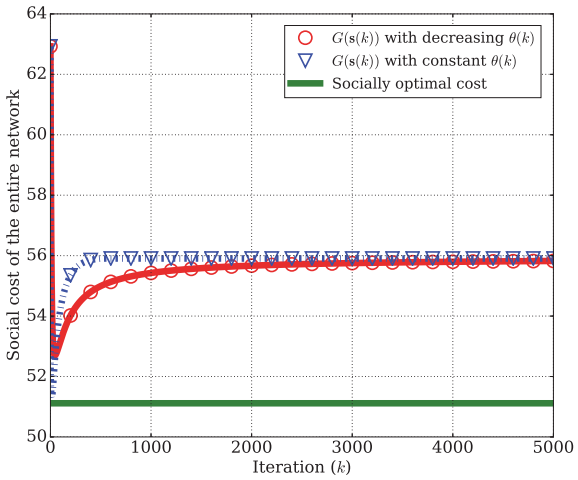



Fig. 5. Evolution of social cost without a channel congestion pricing.

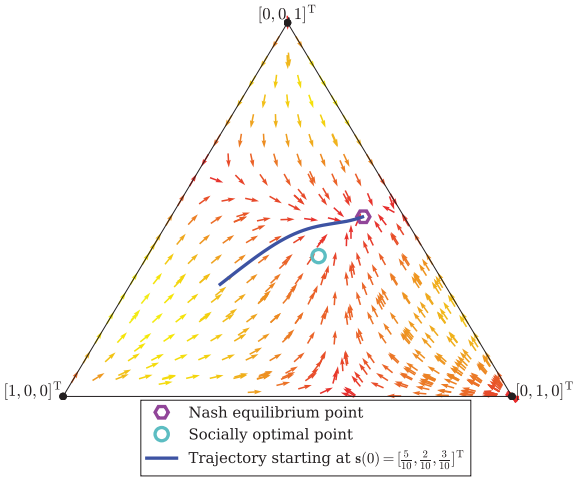



Fig. 6. Evolutionary stability without a channel congestion pricing.

also see that the evolutionary dynamics with the dynamic learning rate vanishing over time converges more slowly than that with the constant learning rate. The reason is that the learning rate is shrinking along with the evolution and the frequency of access strategy adaptation of the vehicular nodes controlled by the learning rate becomes small after a certain number of decision-making iterations.

Furthermore, we examine the network-wide performance of the Nash equilibrium solution and compare it with the ideal socially optimal performance, as shown in Fig. 5. As expected, due to the inherent inefficiency of the Nash equilibrium, it cannot minimize the social cost utility of the vehicular nodes and there exists a gap between the performance of the Nash equilibrium-based solution and the optimal performance. Fig. 6 shows the direction field and phase trajectory of the evolutionary dynamics with the dynamic learning rate as well as its Nash equilibrium and the optimal solution in the plane of the corresponding 2-dimensional simplex $\Delta^2$. This figure also confirms that without a pricing mechanism the Nash equilibrium point does not coincide with the socially optimal state. In Fig. 6, each arrow indicates the direction of evolution of the vehicular population, and its color shade indicates the growth rate of the population shares. From Fig. 6, we can see that, as confirmed by our derived theorem, the Nash equilibrium, i.e., an interior evolutionary equilibrium, is an evolutionary stable strategy, whose asymptotic stability can also be guaranteed in the interior region of the simplex (i.e., all the potential phase trajectories in the interior simplex space will evolve to the Nash equilibrium, as indicated by the direction field in Fig. 6).

### 6.2.2 *Evolutionary dynamics with delayed channel pricing*

We investigate the evolutionary dynamics with the proposed delayed dynamic channel congestion pricing mechanism. In this experiment, we set the epoch size, $W$, of adjusting the channel prices as $W = 200$, and different types of the learning rates are set as $\theta(k) = 10/(100 + k)$ for $\forall k$ and $\theta = 0.1$, respectively. The evolutions of the population shares, the population payoff and the channel prices are illustrated in Figs. 7, 8 and 9, respectively. From these figures, we can find that the introduction of a finite delay $W$ into the channel pricing affects the convergence speed of the evolutionary dynamics. However, as expected, the vehicular nodes can still achieve an evolutionary equilibrium

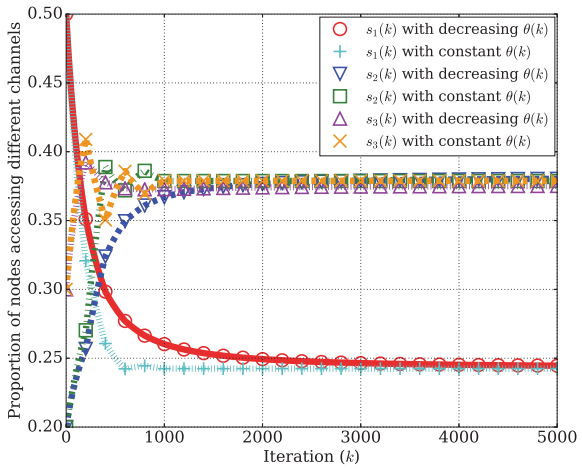



Fig. 7. Evolution of population shares with a delayed pricing.

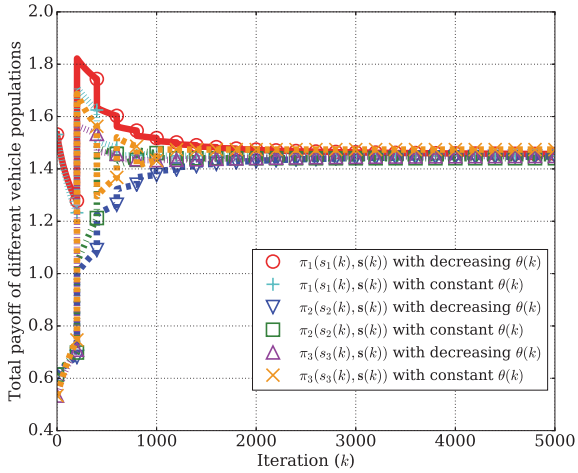



Fig. 8. Evolution of populations' payoffs with a delayed pricing.

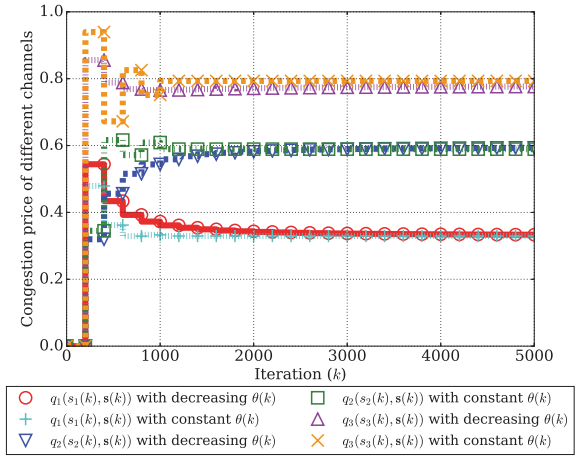



Fig. 9. Evolution of channel prices with a delay $W = 200$.

as shown in Fig. 7, which is also the Nash equilibrium of the evolutionary game, since all the cost utilities experienced by the different vehicular nodes are equal to the average level (See Fig. 8). From Fig. 9, it can be seen that the curve of each dynamic channel congestion price evolves as a piecewise

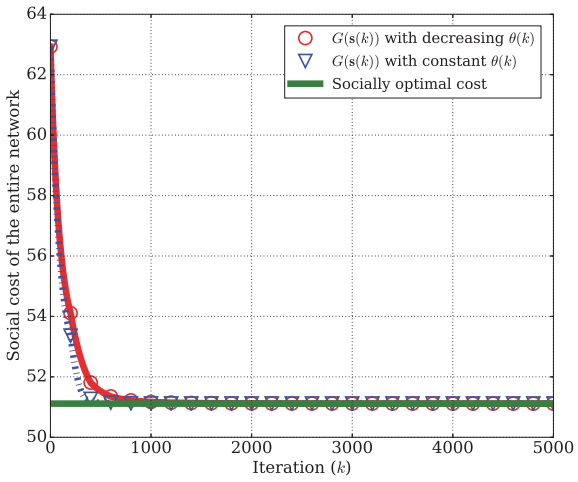



Fig. 10. Evolution of social cost with a delayed pricing.

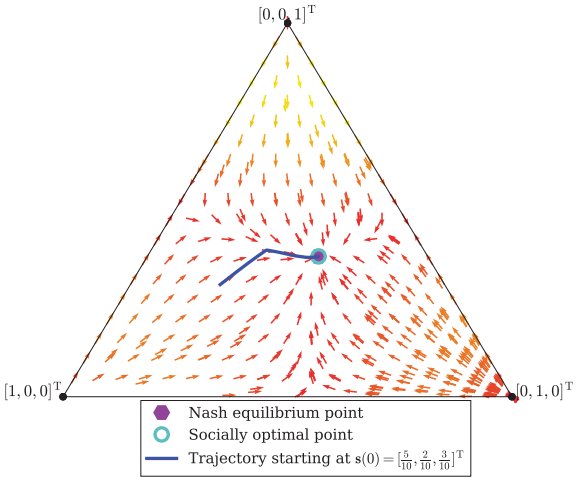



Fig. 11. Evolutionary stability with a delayed pricing.

function with $W = 200$, while the price curve under the constant learning rate can converge faster than that under the vanishing learning rate. Fig. 9 also indicates that the vehicular nodes contending for accessing a more congested channel (i.e., the population share $s_3(k)$ is larger in Fig. 7), e.g., channel 3, will be charged with a higher channel price $q_3(s_3(k), \boldsymbol{s}(k))$.

Fig. 10 compares the performance of the Nash equilibrium-based channel access solution obtained from the evolutionary dynamics with the proposed delayed channel pricing mechanism with the social optimum. It confirms that the proposed dynamics with the delayed pricing can converge to the optimal solution. Fig. 11 also reveals that the Nash equilibrium obtained by the proposed dynamics with the vanishing learning rate well coincides with the socially optimal point. Additionally, the figure demonstrates that the evolutionary stability of the proposed dynamics with the delayed pricing can be guaranteed. That is, the vehicular nodes with the proposed dynamics can always converge to the optimal solution, i.e., the interior evolutionary equilibrium, from any initialization within the interior space of the simplex or can return to the optimal state even with any small local perturbation. The interior evolutionary

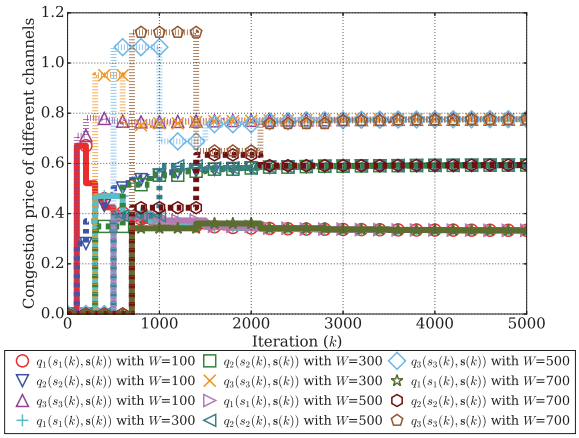



Fig. 12. Evolution of channel prices with different delays.

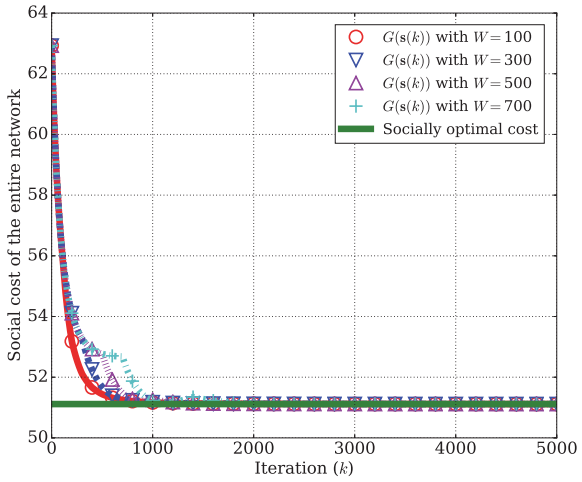



Fig. 13. Evolution of social cost with different delays.

equilibrium point is indeed the Nash equilibrium, which is also the evolutionary stable strategy that has the asymptotic stability.

### 6.2.3 Impact of pricing delay

To investigate the impact of different pricing delays, we implement the evolutionary dynamics with the dynamic learning rate $\theta(k) = 30/(300 + k)$ for all $k$, and then compare the numerical results under different delay parameters $W \in \{100, 300, 500, 700\}$. Fig. 12 shows the channel price curves corresponding to different $W$. As expected, the larger the pricing delay is, the slower the convergence of the evolutionary dynamics becomes. Nevertheless, when the number of the evolution epoch becomes very large, i.e., much larger than the finite $W$, the vehicular nodes can still converge to the equilibrium state. The impact of different pricing delays on the convergence is also shown in Fig. 13. This figure reveals that the proposed dynamics even with different finite pricing delays under the vanishing learning rate can converge to the optimal steady performance as long as the number of the epochs is sufficiently large.

### 6.2.4 Impact of time-varying environment

It is interesting to examine the adaptability of the proposed dynamics under the dynamic learning rate and the constant

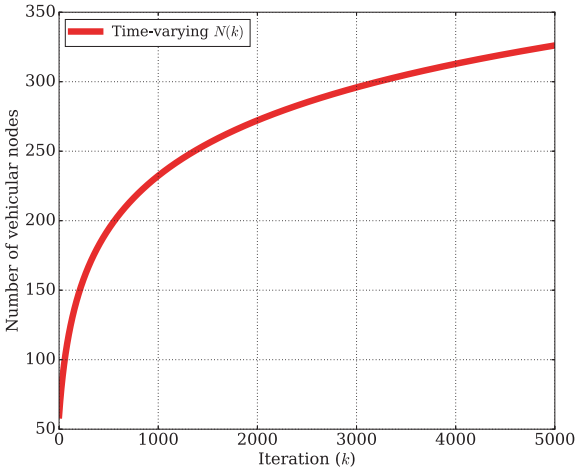



Fig. 14. Time-varying population size.

learning rate, respectively. Hence, we further conduct an experiment, in which we consider the time-varying environment, i.e., the time-varying population size $N(k)$, as shown in Fig. 14. In the dynamic context, we consider that the vehicular population size is gradually increasing along with the evolution process and the grow rate of the vehicular population is gradually decreasing (See Fig. 14). The channel pricing delay is fixed at $W = 200$. Besides, we use the dynamic learning rate $\theta(k) = 0.2/(400 + k)$ for all $k$ and $\theta = 0.0005$.

Fig. 15 shows the evolution of the population state under these two learning rates in the dynamic environment. As we can expect, the evolutionary dynamics with the vanishing learning rate cannot well accommodate the time-varying environment, while the vehicular nodes with the constant learning rate can converge much faster to the equilibrium state. This fact is also confirmed in Fig. 16, which compares that adaptation of the channel prices under the two learning rates. Furthermore, from Fig. 17, we can also observe that when the iteration number becomes large, the constant learning rate can induce the vehicular nodes to closely approach the socially optimal solution. Since the dynamic learning rate becomes smaller and smaller along with the iteration, it cannot capture the time-dependent dynamics of the environment. Consequently, the vanishing learning rate cannot efficiently adapt the access strategies of the vehicular nodes. Clearly, combining all the numerical results in this figures, we can see that there exists a trade-off in the selection of a learning rate in terms of the system adaptability and its optimal performance.

### 6.2.5 Performance comparison

To demonstrate the advantages of the proposed dynamic evolutionary game-theoretic method with the adaptive channel pricing mechanism (marked as ''DEG''), we compare it with a well-known distributed reinforcement learning scheme (marked as ''DRL'') proposed in the related work [18]. In [18], the authors proposed the distributed reinforcement learning scheme (also named Q-learning approach) for mobile users in heterogeneous accessing networks to achieve the Nash-equilibrium solution in an evolutionary game. It is also worth pointing out that the rein-

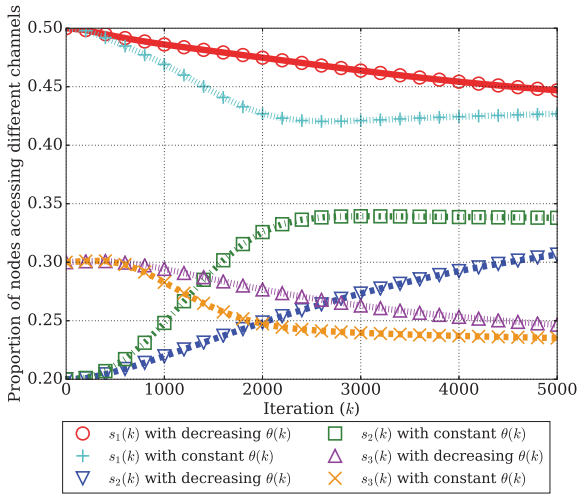



Fig. 15. Evolution of population shares in a time-varying environment.

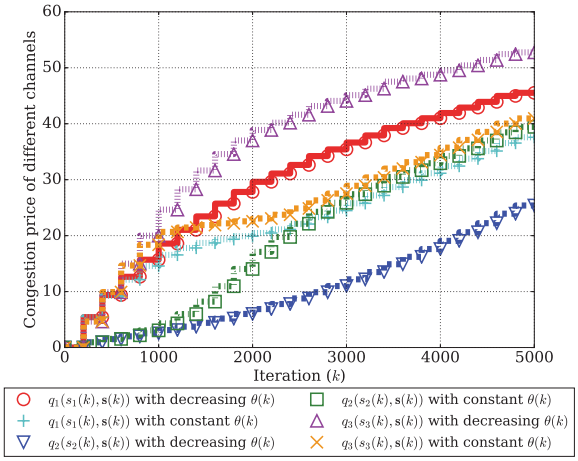



Fig. 16. Evolution of channel prices with a delay $W = 200$ in a time-varying environment.

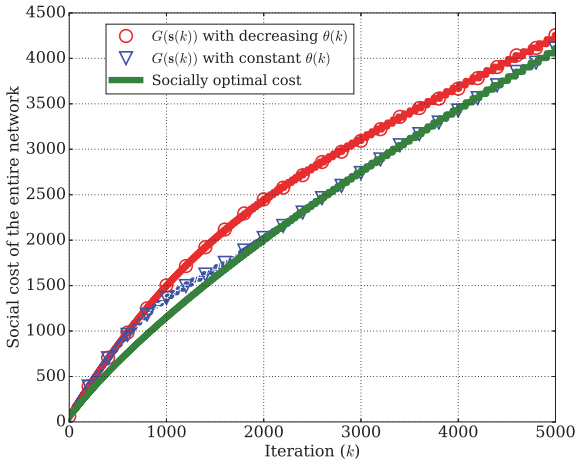



Fig. 17. Evolution of social cost in a time-varying environment.

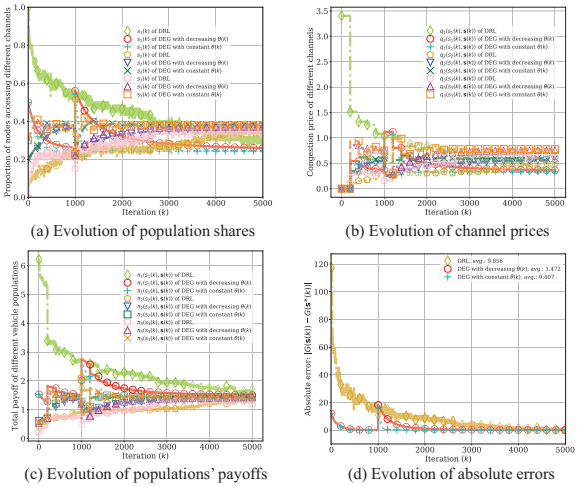
forcement learning mechanism or its variant is also widely used in other evolutionary game-theoretic works focusing on the issue of radio spectrum access such as [40] and [24]. Specifically, we implement the distributed reinforcement learning scheme in a multi-agent setting. For any vehicular node $j$, it needs to maintain a Q-value associated with a channel $i$ in each epoch $k$, denoted by $Q_{i,j}(k)$. Let $p \in (0,1)$ denote a small probability with which the vehicular node $j$ performs a random exploration, i.e., determining an access channel by random selection. In the reinforcement learning approach, the vehicular node $j$ can decide its access channel $a_j(k)$ at $k$ by the stochastic exploration rule or by the exploitation rule: i) if a uniform random number generated at $k$, $\psi(k) = rand()$, satisfies $\psi(k) \leq p$, then $j$ determines its access channel by uniformly and randomly selecting a channel from $\mathcal{M}$ as $a_j(k)$; ii) otherwise, $j$ decides $a_j(k) = \text{argmax}_{\forall i}\{Q_{i,j}(k)\}$ as its targeted access channel. As long as a new payoff associated with a channel $i$ is observed, $\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$, the individual Q-value can be updated by

$$Q_{i,j}(k+1) = (1 - \lambda I_{\{a_j(k)=i\}})Q_{i,j}(k) + \lambda I_{\{a_j(k)=i\}}\left[\frac{1}{\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))} + \beta \max_{\forall i}\{Q_{i,j}(k)\}\right], \tag{32}$$

where $\lambda$ and $\beta$ are the learning rate and the discount factor for the reinforcement learning, respectively. $I_{\{a_j(k)=i\}}$ is an indicator function which is equal to 1 only if $a_j(k) = i$, otherwise 0. Recalling that $\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$ is a cost-type payoff, we transform it into a benefit-type payoff as $1/\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$ that can be used in the payoff maximization formulation as in (32).

To give an insight into the performance in terms of the system convergence, stability and adaptability, we carry out the performance comparison in three different simulation environments, which include a static environment where the population size $N$ is fixed, a dynamic environment with time-varying population size as well as a stochastic environment where the population size is dynamically and stochastically changing. It should be noted that the parameters of both the compared methods need to be appropriately set in different application situations in order to guarantee the system convergence and good performance. In addition, to compare the system optimality, we calculate the time series of the absolute error between the actual social cost obtained by a scheme, $G(\boldsymbol{s}(k))$, and the ideal optimal cost, $G(\boldsymbol{s}^*(k))$, by $|G(\boldsymbol{s}(k)) - G(\boldsymbol{s}^*(k))|$. The corresponding average result can also be evaluated by $\sum_{k=1}^{T}|G(\boldsymbol{s}(k)) - G(\boldsymbol{s}^*(k))|/T$ where $T$ is the total epoch number. Numerical results are as follows.

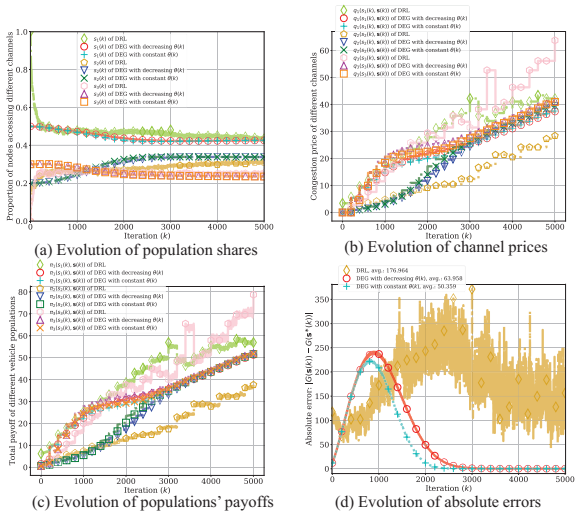
1) *Performance comparison in a static environment*: To simulate a static environment, we fix the population size $N$ at $N = 60$ and set the channel pricing delay as $W = 200$. Besides, to show the evolutionary stability of our proposed method, we introduce a random strategic mutation for 30% players of the whole population at the iteration $k = 1000$, i.e., uniformly and randomly sampling 30% players from the whole population and then changing their current strategies. Such a random strategic mutation can be treated as a perturbation for the system. The resulting evolutions of the population shares, the channel prices and the payoffs of the different subpopulations are shown in Figs. 18(a), (b) and (c), respectively. It can be seen from Fig. 18(a) that by using the proposed evolutionary dynamics, the




(a) Evolution of population shares (b) Evolution of channel prices

(c) Evolution of populations' payoffs (d) Evolution of absolute errors

Fig. 18. Performance comparison in a static environment with a constant population size $N = 60$ while a random strategic mutation is implemented for $30\%$ of the global population at $k = 1000$. Adopted parameter settings include: $W = 200$, $\theta(k) = 30/(300 + k)$ for the dynamic learning rate, $\theta(k) = 0.1$ for the constant learning rate, $\lambda = 0.05$, $p = 0.01$, $\beta = 0.05$.




(a) Evolution of population shares (b) Evolution of channel prices

(c) Evolution of populations' payoffs (d) Evolution of absolute errors

Fig. 19. Performance comparison in a dynamic environment with increasing population size as given in Fig. 14. Adopted parameter settings include: $W = 200$, $\theta(k) = 2/(4 \times 10^3 + k)$ for the dynamic learning rate, $\theta(k) = 5 \times 10^{-4}$ for the constant learning rate, $\lambda = 0.005$, $p = 0.01$, $\beta = 0.005$.

system can converge to an equilibrium point (also a Nash equilibrium) with the population shares associated with the 2nd and the 3rd channels evolving to almost the same level around $0.37$ while that of the 1st channel converging to around $0.25$. Another important fact can also be observed that even though a random perturbation is imposed on the evolving population, our method can effectively restore to the equilibrium point, indicating that the evolutionary equilibrium point obtained is exactly the ESS and that the system stability can be well guaranteed. In contrast, with the distributed reinforcement learning scheme, the system can only converge to a neighbor of the equilibrium point within the given iteration epochs. From Figs. 18(b) and (c), we can also see that after the perturbation at $k = 1000$, our proposed method can better induce the channel prices to converge to an equilibrium state and the payoffs of the channels to the same level, which indicates that the system arrives at the Nash equilibrium since all the players can gain nothing by unilaterally changing their own strategies when all the channel payoffs are the same. Additionally, Fig. 18(d) illustrates the convergence of the absolute error between the actual social cost and the ideal optimal cost. Compared to the distributed reinforcement learning with the average absolute error of 9.856, the absolute error of our method converges to a much lower level with the average values of 1.472 and 0.407 for the dynamic learning rate and for the constant learning rate, respectively. This figure shows that our method can faster converge to around zero even with the strategic perturbation occurring at $k = 1000$, indicating that the attained Nash equilibrium is socially efficient and evolutionarily stable.

2) *Performance comparison in a dynamic environment*: In a dynamic environment, we increase the population size $N$ from 60 to around 325 over time, as shown in Fig. 14. From Fig. 19, by comparison, it is obvious that our method can adapt to the dynamically changing environment in a more effective manner than the distributed reinforcement learning scheme. In Fig. 19(a), it is seen that the population shares in a equilibrium state in such a dynamic setting are different from each other. Besides, Figs. 19(b) and (c) show that our proposed evolutionary dynamics with the dynamic learning rate and the constant learning rate can better guarantee the convergence of both the channel prices and the payoffs while the compared scheme cannot stably converge within the given epochs. This fact is more obvious from Fig. 19(d). As shown in Fig. 19(d), the absolute error of the proposed evolutionary dynamics based on the constant learning rate can faster converge than that based on the dynamic learning rate, which confirms that the constant learning rate achieves better adaptability. Nevertheless, the absolute error of the proposed method implemented with both the learning rates can finally almost converge to zero (the average result of the dynamic learning rate is 63.958 and that of the constant learning rate is 50.359), while the absolute error of the compared scheme with the average result of 176.964 is shown to be fluctuating over time and cannot converge to the neighbor of zero within the given epochs.

3) *Performance comparison in a stochastic environment*: Moreover, we compare our method with the distributed reinforcement learning scheme in a stochastic environment where the population size $N$ is assumed to be time-varying and follows a Poisson point process characterized by the average rate $\mathbb{E}[N(k)] = 200$. The evolution of $N(k)$ over epochs $k$ and its distribution in the simulation experiment are shown in Fig. 20. It can be observed from Figs. 21(a), (b) and (c) that the evolutions of the population shares, the channel prices and the channel payoffs with our proposed method are less sensitive to the randomness in the time-varying environment than those with the compared scheme. Furthermore, Fig. 21(d) shows that the average of the absolute error of our method is 20.236 for the dynamic learning rate and 19.068 for the constant learning rate, which is much less than that of the compared scheme (the average of the absolute error of the compared scheme is 183.177).

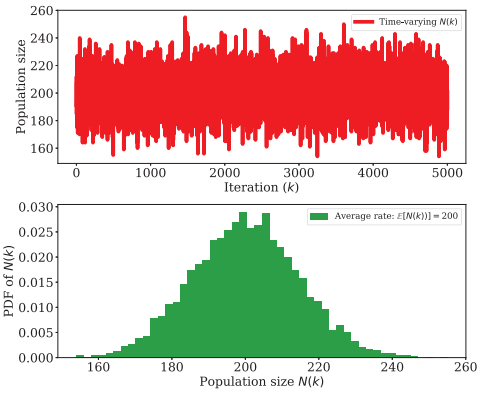



Fig. 20. Stochastic time-varying population size which follows a Poisson point process, i.e., $N(k) \sim \text{Pois}(\mathbb{E}[N(k)])$ with the average rate $\mathbb{E}[N(k)] = 200$.

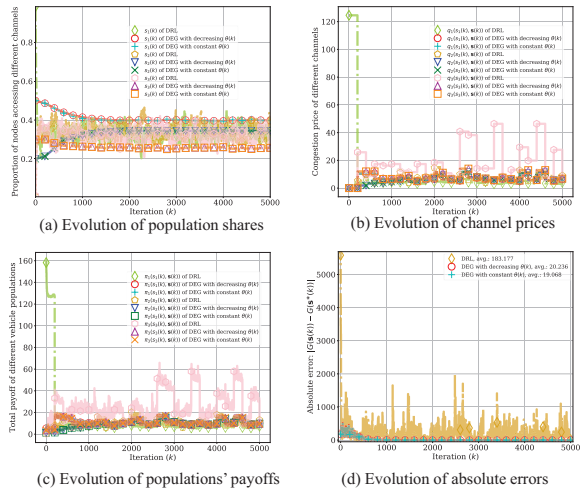



Fig. 21. Performance comparison in a dynamic and stochastic environment with stochastic population size as given in Fig. 20. Adopted parameter settings include: $W = 200$, $\theta(k) = 2 \times 10^3/(4 \times 10^3 + k)$ for the dynamic learning rate, $\theta(k) = 0.5$ for the constant learning rate, $\lambda = 0.05$, $p = 0.01$, $\beta = 0.05$.

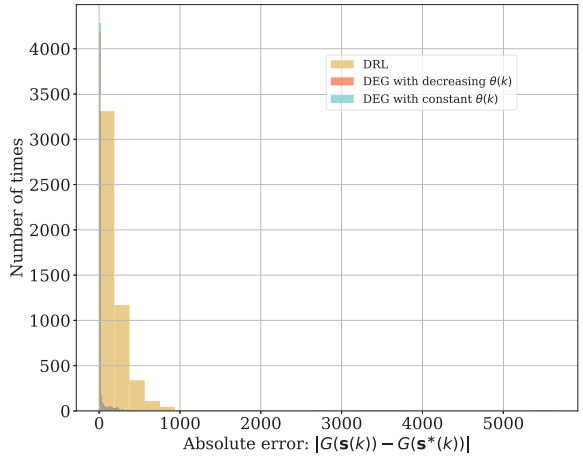



Fig. 22. The distribution of the absolute error between the actual social cost $G(\boldsymbol{s}(k))$ and the ideal optimal cost $G(\boldsymbol{s}^*(k))$ in a dynamic and stochastic environment.

Another fact can also be confirmed from Fig. 21(d) that our method can drive the system to faster adapt to the fluctuating environment since the corresponding absolute error associated with either the dynamic learning rate or the constant learning rate can stably converge to around zero, while the absolute error of the distributed reinforcement learning scheme fluctuates above zero. Besides, we also illustrate the distributions (the histograms) of the resulting absolute errors in Fig. 22, which clearly shows that the proposed evolutionary dynamics (with both types of the learning rate) has much more absolute errors distributed near zero than the compared scheme does. This indicates that our method can achieves better system-wide performance in this stochastic environment by adapting the system to converge to an equilibrium closer to the ideal optimal point.

## 7 Conclusion and Future Work

We have formulated a discretized evolutionary dynamics combined with a delayed channel pricing mechanism to optimize the channel access of competitive vehicular nodes in a cognitive vehicular network. We have established the theoretical connections between the evolutionary equilibrium, the Nash equilibrium, and the evolutionary stable strategy of the formulated evolutionary game. Besides, we analytically derive the bound of the difference between the proposed evolutionary dynamics with the delayed pricing and that without a delay in pricing, proving that the asymptotic stability of the proposed dynamics can be guaranteed. We also obtain the theoretical bound of the difference between the steady performance of the proposed dynamics and the ideal optimal state under the shrinking and the constant learning rates, showing that the shrinking learning rate can induce the vehicular nodes to asymptotically converge to the Nash equilibrium which is also the optimum point. In contrast, we can only prove that the proposed dynamics under the constant learning rate potentially approaches a neighborhood of the optimal point. Nonetheless, applying the constant learning rate can better adapt the cognitive vehicular network to a dynamic environment, highlighting an important trade-off between the system adaptability and optimality. In the future research, it is expected to incorporate the impacts of not only the dynamics of the channel competition from the perspective of the communication system but also the temporal-spatial distribution of the vehicles from the perspective of the transportation system.

## Acknowledgments

This research was supported in part by the National Natural Science Foundation of China under Grant Nos. 61672082 and 61822101, Beijing Municipal Natural Science Foundation No.4181002, the scholarship under the State Scholarship Fund (No. 201706020112), Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Asa Briggs Visiting Fellowship from University of Sussex, Royal Society-Newton Mobility Grant (IE160920) and The Engineering, and Physical Sciences Research Council (EPSRC) (EP/P025862/1).

## REFERENCES


[1] J. B. Kenney, "Dedicated short-range communications (dsrc) standards in the united states," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1162–1182, July 2011.

[2] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE personal communications*, vol. 6, no. 4, pp. 13–18, 1999.

[3] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer networks*, vol. 50, no. 13, pp. 2127–2159, 2006.

[4] M. M. Buddhikot, "Understanding dynamic spectrum access: Models,taxonomy and challenges," in *Proceedings of the 2007 2Nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 649–663. [Online]. Available: http://dx.doi.org/10.1109/DYSPAN.2007.88

[5] M. D. Felice, R. Doost-Mohammady, K. R. Chowdhury, and L. Bononi, "Smart radios for smart vehicles: Cognitive vehicular networks," *IEEE Vehicular Technology Magazine*, vol. 7, no. 2, pp. 26–33, June 2012.

[6] J. W. Weibull, *Evolutionary game theory*. MIT press, 1997.

[7] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bulletin of the American Mathematical Society*, vol. 40, no. 4, pp. 479–519, 2003.

[8] P. Dubey, "Nash equilibria of market games: Finiteness and inefficiency," *Journal of Economic Theory*, vol. 22, no. 2, pp. 363 – 376, 1980. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0022053180900484

[9] ——, "Inefficiency of nash equilibria," *Mathematics of Operations Research*, vol. 11, no. 1, pp. 1–8, 1986. [Online]. Available: https://doi.org/10.1287/moor.11.1.1

[10] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 192–202, Jan 2008.

[11] A. C. Pigou, *The economics of welfare*. Palgrave Macmillan, 2013.

[12] H. Yang and H.-J. Huang, "Principle of marginal-cost pricing: how does it work in a general road network?" *Transportation Research Part A: Policy and Practice*, vol. 32, no. 1, pp. 45 – 54, 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0965856497000189

[13] D. Niyato, E. Hossain, and P. Wang, "Optimal channel access management with qos support for cognitive vehicular networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 4, pp. 573–591, 2011.

[14] R. Alhamad, H. Wang, and Y. D. Yao, "Cooperative spectrum sensing with random access reporting channels in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 7249–7261, Aug 2017.

[15] E. Hossain, D. Niyato, and Z. Han, *Dynamic spectrum access and management in cognitive radio networks*. Cambridge university press, 2009.

[16] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y. D. Yao, "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1380–1391, 2012.

[17] Z. Dai, Z. Wang, and V. W. S. Wong, "An overlapping coalitional game for cooperative spectrum sensing and access in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8400–8413, Oct 2016.

[18] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2008–2017, 2009.

[19] ——, "A noncooperative game-theoretic framework for radio resource management in 4g heterogeneous wireless access networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 3, pp. 332–345, March 2008.

[20] M. Maskery, V. Krishnamurthy, and Q. Zhao, "Decentralized dynamic spectrum access for cognitive radios: cooperative design of a non-cooperative game," *IEEE Transactions on Communications*, vol. 57, no. 2, pp. 459–469, February 2009.

[21] A. Attar, M. R. Nakhai, and A. H. Aghvami, "Cognitive radio game for secondary spectrum access problem," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 2121–2131, April 2009.

[22] S. Wang, H. Liu, P. H. Gomes, and B. Krishnamachari, "Deep Reinforcement Learning for Dynamic Multichannel Access in Wireless Networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 2, pp. 257–265, 2018.

[23] M. Yan, G. Feng, J. Zhou, and S. Qin, "Smart Multi-RAT Access Based on Multiagent Reinforcement Learning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4539–4551, 2018.

[24] X. Chen and J. Huang, "Evolutionarily Stable Spectrum Access," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1281–1293, 2013.

[25] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, 2015.

[26] H. Cui, Y. Wang, Q. Guan, and H. Zhang, "Distributed interference-aware cooperative mac based on stackelberg pricing game," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4124–4134, Sept 2015.

[27] Q. Zhang, B. Fu, Z. Feng, and W. Li, "Utility-maximized two-level game-theoretic approach for bandwidth allocation in heterogeneous radio access networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 844–854, Jan 2017.

[28] K. Zhu, E. Hossain, and D. Niyato, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1843–1856, Aug 2014.

[29] X. Kang, R. Zhang, and M. Motani, "Price-based resource allocation for spectrum-sharing femtocell networks: A stackelberg game approach," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, pp. 538–549, April 2012.

[30] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: a dynamic game approach," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2651–2660, July 2008.

[31] D. Niyato, E. Hossain, and Z. Han, "Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: A game-theoretic modeling approach," *IEEE Transactions on Mobile Computing*, vol. 8, no. 8, pp. 1009–1022, Aug 2009.

[32] D. Niyato and E. Hossain, "Spectrum trading in cognitive radio networks: A market-equilibrium-based approach," *IEEE Wireless Communications*, vol. 15, no. 6, pp. 71–80, December 2008.

[33] ——, "Market-equilibrium, competitive, and cooperative pricing for spectrum sharing in cognitive radio networks: Analysis and comparison," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4273–4283, November 2008.

[34] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y. D. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 2, pp. 180–194, 2012.

[35] R. Chen, Z. Sheng, Z. Zhong, M. Ni, V. C. M. Leung, D. G. Michelson, and M. Hu, "Connectivity analysis for cooperative vehicular ad hoc networks under nakagami fading channel," *IEEE Communications Letters*, vol. 18, no. 10, pp. 1787–1790, Oct 2014.

[36] T. H. Luan, X. S. Shen, and F. Bai, "Integrity-oriented content transmission in highway vehicular ad hoc networks," in *2013 Proceedings IEEE INFOCOM*, April 2013, pp. 2562–2570.

[37] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated arq over wireless links," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1746–1755, Sept 2004.

[38] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge University Press Cambridge, 2007, vol. 1.

[39] V. S. Borkar *et al.*, "Stochastic approximation," *Cambridge Books*, 2008.

[40] H. Li, "Multiagent q-learning for aloha-like spectrum access in cognitive radio systems," *EURASIP J. Wirel. Commun. Netw.*, vol. 2010, pp. 56:1–56:13, Apr. 2010. [Online]. Available: http://dx.doi.org/10.1155/2010/876216

**Daxin Tian** [M'13, SM'16] is an associate professor in the School of Transportation Science and Engineering, Beihang University, Beijing, China. His current research interests include mobile computing, intelligent transportation systems, vehicular ad hoc networks, and swarm intelligent.

**Jianshan Zhou** received the B.Sc. and M.Sc. degrees in traffic information engineering and control in 2013 and 2016, respectively. He is currently working towards the Ph.D. degree with the School of Transportation Science and Engineering, Beihang University, Beijing, China. His current research interests are focused on wireless communication, artificial intelligent system, and intelligent transportation systems.

**Yunpeng Wang** is a professor in the School of Transportation Science and Engineering, Beihang University, Beijing, China. His current research interests include intelligent transportation systems, traffic safety, and vehicle infrastructure integration.

**Zhengguo Sheng** is currently a Lecturer with the Department of Engineering and Design, University of Sussex, U.K. He has authored over 50 international conference and journal papers. His current research interests cover IoT/M2M, vehicular communications, and edge/cloud computing.

**Xuting Duan** is currently an assistant professor with the School of Transportation Science and Engineering, Beihang University, Beijing, China. His current research interests are focused on vehicular ad hoc networks.

**Victor C. M. Leung** [S'75, M'89, SM'97, F'03] received the B.A.Sc. (Hons.) degree in electrical engineering from the University of British Columbia (UBC) in 1977, and was awarded the APEBC Gold Medal as the head of the graduating class in the Faculty of Applied Science. He attended graduate school at UBC on a Canadian Natural Sciences and Engineering Research Council Postgraduate Scholarship and received the Ph.D. degree in electrical engineering in 1982.

From 1981 to 1987, Dr. Leung was a Senior Member of Technical Staff and satellite system specialist at MPR Teltech Ltd., Canada. In 1988, he was a Lecturer in the Department of Electronics at the Chinese University of Hong Kong. He returned to UBC as a faculty member in 1989, and currently holds the positions of Professor and TELUS Mobility Research Chair in Advanced Telecommunications Engineering in the Department of Electrical and Computer Engineering. Dr. Leung has co-authored more than 1000 journal/conference papers, 37 book chapters, and co-edited 12 book titles. Several of his papers had been selected for best paper awards. His research interests are in the broad areas of wireless networks and mobile systems.

Dr. Leung is a registered Professional Engineer in the Province of British Columbia, Canada. He is a Fellow of IEEE, the Royal Society of Canada, the Engineering Institute of Canada, and the Canadian Academy of Engineering. He was a Distinguished Lecturer of the IEEE Communications Society. He is serving on the editorial boards of the IEEE Wireless Communications Letters, IEEE Transactions on Green Communications and Networking, IEEE Access, Computer Communications, and several other journals, and has previously served on the editorial boards of the IEEE Journal on Selected Areas in Communications  Wireless Communications Series and Series on Green Communications and Networking, IEEE Transactions on Wireless Communications, IEEE Transactions on Vehicular Technology, IEEE Transactions on Computers, and Journal of Communications and Networks. He has guest-edited many journal special issues, and provided leadership to the organizing committees and technical program committees of numerous conferences and workshops. He received the IEEE Vancouver Section Centennial Award and 2011 UBC Killam Research Prize. He is the recipient of the 2017 Canadian Award for Telecommunications Research. He is a co-author of the paper that has won the 2017 IEEE ComSoc Fred W. Ellersick Prize.

# Supplementary Materials for Channel Access Optimization with Adaptive Congestion Pricing for Cognitive Vehicular Networks: An Evolutionary Game Approach

Daxin Tian, *Senior Member, IEEE,* Jianshan Zhou, Yunpeng Wang, Zhengguo Sheng, Xuting Duan, and Victor C.M. Leung, *Fellow, IEEE*

## 1 Proof of Theorem 1

The optimization objective of (9), denoted by $F(\boldsymbol{s}) = \sum_{i=1}^{M}\int_0^{s_i} f_i(x,\boldsymbol{s})dx$, is a continuous version of Rosenthal's potential function [1]. Therefore, the formulated game is a potential game, for which there exists at least one Nash equilibrium. Furthermore, this result also guarantees that a minimum point of the potential function $F(\boldsymbol{s})$ is a Nash equilibrium. Notice that the optimization objective function in (9) is strictly convex and the set of $\boldsymbol{s}$ is compact. The model above is indeed a convex optimization problem, which can ensure that its local minimum arises as its global minimizer and the global minimizer always and uniquely exist. In other words, the existence and uniqueness of the Nash equilibrium is guaranteed. To obtain the Nash equilibrium comes down to solving the convex optimization.

## 2 Proof of Theorem 2

The definitions of $\boldsymbol{s}^{\mathrm{NE}}$ and $\boldsymbol{s}^*$ obviously make $1 \leq \theta$ held. So we only need to prove $\theta \leq \frac{\lambda}{1-\mu}$. Furthermore, based on Lemmas 1 and 2, we can directly obtain $\sum_{i=1}^{M} s_i^{\mathrm{NE}} f_i\left(s_i^{\mathrm{NE}}, \boldsymbol{s}^{\mathrm{NE}}\right) \leq \lambda \sum_{i=1}^{M} s_i^* f_i\left(s_i^*, \boldsymbol{s}^*\right) + \mu \sum_{i=1}^{M} s_i^{\mathrm{NE}} f_i\left(s_i^{\mathrm{NE}}, \boldsymbol{s}^{\mathrm{NE}}\right)$, which implies

$$\frac{\sum_{i=1}^{M} s_i^{\mathrm{NE}} f_i\left(s_i^{\mathrm{NE}}, \boldsymbol{s}^{\mathrm{NE}}\right)}{\sum_{i=1}^{M} s_i^* f_i\left(s_i^*, \boldsymbol{s}^*\right)} \leq \frac{\lambda}{1-\mu}. \tag{S.1}$$

At this point, Theorem 2 is proven.

## 3 Proof of Theorem 3

Notice that both the models (11) and (15) follow the same constraints. Thus, they have the same feasible domain, implying that the feasible solution of (11) is also the feasible solution of (15) and vice versa. Moreover, they have strict convex objective functions over the same convex set, which means that i) their local minimizers are also their global minimizers and ii) the global minimizers are unique.

Now, we let $\boldsymbol{s}^*$ and $\boldsymbol{z}^{\mathrm{NE}}$ be the unique minimizers of (11) and (15), respectively. We only need to show that $\boldsymbol{s}^*$ and $\boldsymbol{z}^{\mathrm{NE}}$ are identical to each other to complete the proof. Denote $F_\pi(\boldsymbol{s}) = \sum_{i=1}^{M}\int_0^{s_i} \pi_i(x,\boldsymbol{s})dx$. It is obvious that $\bigtriangledown_{\boldsymbol{s}} F_\pi(\boldsymbol{s}) = \frac{1}{N}\bigtriangledown_{\boldsymbol{s}} G(\boldsymbol{s})$. Based on the Karush-Kuhn-Tucker (KKT) conditions, given the socially optimal point $\boldsymbol{s}^*$ of (11) (it is the feasible points of (11) and (15)), its associated KKT multipliers $w_l^* \in \mathbb{R}_{\geq 0}$ and $v^* \in \mathbb{R}$ must exist, such that $\bigtriangledown_{\boldsymbol{s}} G\left(\boldsymbol{s}^*\right) = \sum_{l\in\mathcal{A}(\boldsymbol{s}^*)} w_l^* \bigtriangledown_{\boldsymbol{s}} g_l\left(\boldsymbol{s}^*\right) + v^* \bigtriangledown_{\boldsymbol{s}} h\left(\boldsymbol{s}^*\right)$, where $\mathcal{A}(\boldsymbol{s}^*)$ is the active set of the inequalities associated with $\boldsymbol{s}^*$, $g_l\left(\boldsymbol{s}\right) = s_l$ and $h(\boldsymbol{s}) = 1 - \sum_{l=1}^{M} s_l$. Therefore, we can also correspondingly obtain $\bigtriangledown_{\boldsymbol{s}} F_\pi\left(\boldsymbol{s}^*\right) = \frac{1}{N}\bigtriangledown_{\boldsymbol{s}} G\left(\boldsymbol{s}^*\right) = \sum_{l\in\mathcal{A}(\boldsymbol{s}^*)} \frac{w_l^*}{N} \bigtriangledown_{\boldsymbol{s}} g_l\left(\boldsymbol{s}^*\right) + \frac{v^*}{N} \bigtriangledown_{\boldsymbol{s}} h\left(\boldsymbol{s}^*\right)$. This fact indicates that there also must exist KKT multipliers $\frac{w_l^*}{N} \in \mathbb{R}_{\geq 0}$ and $\frac{v^*}{N} \in \mathbb{R}$ associated with $\boldsymbol{s}^*$ satisfying the KKT conditions of the model (15). Thus, the socially optimal solution $\boldsymbol{s}^*$ must be the unique minimizer of (15), i.e., the Nash equilibrium of the evolutionary game with the channel congestion pricing. By the same means, given the Nash equilibrium of (15), $\boldsymbol{z}^{\mathrm{NE}}$, we can also prove that the KKT conditions of (11) can also be guaranteed by $\boldsymbol{z}^{\mathrm{NE}}$. Recalling the uniqueness of the optimum in (11) and (15), we can prove $\boldsymbol{s}^* = \boldsymbol{z}^{\mathrm{NE}}$.

## 4 Proof of the convexity of the cost-type utility function

For the sake of simplicity, let $\overline{r}_i = \sum_{l=1}^{L} c_{l,i}\mathrm{Prob}_1\left\{c_{l,i}, i\right\} + \sum_{l'=1}^{L} c'_{l'}\mathrm{Prob}_2\left\{c'_{l'}\right\}$. To show the convexity of $f_i(s_i,\boldsymbol{s})$ given in equation (8) withe respect to the population share $s_i$, we examine the property of its second derivative as follows. We first yield the first-order derivative of $f_i(s_i,\boldsymbol{s})$ as

$$f_i'(s_i,\boldsymbol{s}) = \frac{-CN\ln(1-\alpha_i)}{\alpha_i\overline{r}_i(1-\alpha_i)^{Ns_i-1}}. \tag{S.2}$$

Based on (S.2), we then obtain the second derivative by

$$f_i''(s_i,\boldsymbol{s}) = \frac{CN^2\ln^2(1-\alpha_i)}{\alpha_i\overline{r}_i(1-\alpha_i)^{Ns_i-1}}. \tag{S.3}$$

Note that $s_i \in [0,1]$ and $\ln(1-\alpha_i) < 0$ due to the fact $\alpha_i \in (0,1)$. $f_i'(s_i,\boldsymbol{s}) > 0$ as well as $f_i''(s_i,\boldsymbol{s}) > 0$ is always held, so that $f_i(s_i,\boldsymbol{s})$ is a strictly convex function of $s_i$.

## 5 Proof of Theorem 4

First, we would like to show that $\boldsymbol{z}^{\mathrm{NE}}$ is a strict Nash equilibrium. According to the definition of the Nash equilibrium in Definition 1, for any $i \neq i' \in \mathcal{M}$ and $s_i > 0, s_{i'} \geq 0$, $\pi_i\left(s_i^{\mathrm{NE}}, \boldsymbol{z}^{\mathrm{NE}}\right) \leq \pi_{i'}\left(s_{i'}^{\mathrm{NE}}, \boldsymbol{z}^{\mathrm{NE}}\right)$. Suppose that one player switches its strategy from $i$ to $i'$. The strategy deviation results in its new cost $\pi_{i'}\left(s_{i'}^{\mathrm{NE}} + \frac{1}{N}, \boldsymbol{s}'\right)$. Recalling that $\pi_{i'}(s_{i'}, \boldsymbol{s})$ is strictly monotonically increasing with respect to $s_{i'}$ as $\pi'_{i'}(s_{i'}, \boldsymbol{s}) = 2f'_{i'}(s_{i'}, \boldsymbol{s}) + s_i f''_{i'}(s_{i'}, \boldsymbol{s}) > 0$, we can have $\pi_i\left(s_i^{\mathrm{NE}}, \boldsymbol{z}^{\mathrm{NE}}\right) \leq \pi_{i'}\left(s_{i'}^{\mathrm{NE}}, \boldsymbol{z}^{\mathrm{NE}}\right) < \pi_{i'}\left(s_{i'}^{\mathrm{NE}} + \frac{1}{N}, \boldsymbol{s}'\right)$. In other words, the unilateral deviation from $\boldsymbol{z}^{\mathrm{NE}}$ can lead to a strictly higher cost, which indicates that $\boldsymbol{z}^{\mathrm{NE}}$ is exactly a strict Nash equilibrium. Moreover, [2] has shown that any strict Nash equilibrium is also an ESS. Accordingly, $\boldsymbol{z}^{\mathrm{NE}}$ must be an ESS, i.e., $\boldsymbol{z}^{\mathrm{NE}} = \boldsymbol{s}^{\mathrm{ESS}}$.

## 6 Proof of Theorem 5

We first show that $\forall \boldsymbol{s}^{\mathrm{EE}}(t) \in \mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$ always satisfies $\boldsymbol{s}^{\mathrm{EE}}(t) \in \Delta^{M-1}$ for all $t \in \mathbb{R}_{\geq 0}$. On the one side, by initializing (17) at the point $\boldsymbol{s}(0) \in \Delta^{M-1}$, we see $\sum_{i=1}^{M} s_i(0) = 1$. Now, according to the definition of $\mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$, we can always have $d\left(\sum_{i=1}^{M} s_i^{\mathrm{EE}}(t)\right)/dt = \sum_{i=1}^{M}\left(ds_i^{\mathrm{EE}}(t)/dt\right) = 0$ for all $t \in \mathbb{R}_{\geq 0}$, where $s_i^{\mathrm{EE}}(t) \in \boldsymbol{s}^{\mathrm{EE}}(t)$. This means that $\sum_{i=1}^{M} s_i^{\mathrm{EE}}(t) = \sum_{i=1}^{M} s_i(0) = 1 \, \forall \, t \in \mathbb{R}_{\geq 0}$. On the other side, the basic property of the replicator dynamics (17) directly guarantees $\forall s_i^{\mathrm{EE}}(t) \geq 0$ for all $t \in \mathbb{R}_{\geq 0}$ when the initial point is not negative, $\forall s_i(0) \geq 0$. With both the results aforementioned, we can have $\mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}} \subseteq \Delta^{M-1}$. That is to say, all the points in $\mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$ are feasible for the model (11).

Next, it is observed that the definition of $\mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$ coincides with that of the Nash equilibrium given in Definition 1. That is, any $\boldsymbol{s}^{\mathrm{EE}}(t) \in \mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$ is also the Nash equilibrium of (15). Based on Theorem 3, it can be confirmed that $\mathcal{U}_{\mathrm{SO}} = \mathcal{U}_{\mathrm{EE}} \backslash \widetilde{\mathcal{U}}_{\mathrm{EE}}$.

## 7 Proof of Theorem 6

Let $\mathcal{N}(\epsilon) = \left\{\boldsymbol{s} \in \Delta^{M-1} \,|\, \|\boldsymbol{s} - \boldsymbol{s}^*\| < \epsilon, \boldsymbol{s} \neq \boldsymbol{s}^*\right\}$ be the neighborhood of $\boldsymbol{s}^*$. The set of indices of the zero components in $\boldsymbol{s}^*$ can be represented by $\mathcal{A}(\boldsymbol{s}^*) = \{i \in \mathcal{M} \,|\, s_i^* = 0, s_i^* \in \boldsymbol{s}^*\}$. Correspondingly, the set of indices of the nonzero components in $\boldsymbol{s}^*$ can be $\overline{\mathcal{A}}(\boldsymbol{s}^*) = \mathcal{M} \backslash \mathcal{A}(\boldsymbol{s}^*)$.

Now we can conduct the perturbation analysis on $\boldsymbol{s}^*$ by introducing some small perturbations $\{\varepsilon_i(t) > 0 \, \forall i \in \mathcal{M}\}$ that can constitute a disturbance $\boldsymbol{v}(t) = [v_1(t), \ldots, v_M(t)]^{\mathrm{T}}$, in which $v_{i_1}(t) = \varepsilon_{i_1}(t) > 0$ for $i_1 \in \mathcal{A}(\boldsymbol{s}^*)$ while $v_{i_2}(t) = -\varepsilon_{i_2}(t) < 0$ for $i_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)$. Then, we denote a new point deviated from $\boldsymbol{s}^*$ by the disturbance $\boldsymbol{v}(t)$ as $\boldsymbol{s}(t) = \boldsymbol{s}^* + \boldsymbol{v}(t) \in \mathcal{N}(\epsilon)$. To guarantee the feasibility of $\boldsymbol{s}(t)$, the small perturbations are assumed to ensure that $\varepsilon_{i_1}(t) \geq 0$ for $i_1 \in \mathcal{A}(\boldsymbol{s}^*)$, $\varepsilon_{i_2}(t) \in \left[0, s_{i_2}^*\right)$ for $i_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)$, and $\sum_{i_1 \in \mathcal{A}(\boldsymbol{s}^*)} \varepsilon_{i_1}(t) - \sum_{i_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)} \varepsilon_{i_2}(t) = 0$. It can be seen that $\sum_{i=1}^{M} s_i(t) = \sum_{i=1}^{M}\left(s_i^* + v_i(t)\right) = \sum_{i_1 \in \mathcal{A}(\boldsymbol{s}^*)} \varepsilon_{i_1}(t) + \sum_{i_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)}\left(s_{i_2}^* - \varepsilon_{i_2}(t)\right) = \sum_{i_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)} s_{i_2}^* = 1$ always holds. For the sake of simplicity, we use $\pi_i(s_i(t))$ to denote $\pi_i(s_i(t)) = \pi_i(s_i(t), \boldsymbol{s}(t))$. Next, we analyze two cases of $i \in \mathcal{M}$ as follows.

i) For any $i_1 \in \mathcal{A}(\boldsymbol{s}^*)$, following the replicator dynamics (17) we can get $ds_{i_1}(t)/dt = d\varepsilon_{i_i}(t)/dt$ by

$$\frac{d\varepsilon_{i_1}(t)}{dt} = \sigma\varepsilon_{i_1}(t)\left[\sum_{l_1 \in \mathcal{A}(\boldsymbol{s}^*)} \varepsilon_{l_1}(t)\pi_{l_1}\left(\varepsilon_{l_1}(t)\right) + \sum_{l_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)}\left(s_{l_2}^* - \varepsilon_{l_2}(t)\right)\pi_{l_2}\left(s_{l_2}^* - \varepsilon_{l_2}(t)\right) - \pi_{i_1}\left(\varepsilon_{i_1}(t)\right)\right]. \tag{S.4}$$

By linearization, we can neglect the high-order terms of $\varepsilon_{i_1}$, i.e., $\varepsilon_{i_1}(t)\varepsilon_{l_1}(t)$ and $\varepsilon_{i_1}(t)\varepsilon_{l_2}(t)$, and then approximate the above differential equation as

$$\frac{d\varepsilon_{i_1}(t)}{dt} \approx \sigma\varepsilon_{i_1}(t)\left[\begin{array}{l}\displaystyle\sum_{l_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)} s_{l_2}^*\pi_{l_2}\left(s_{l_2}^* - \varepsilon_{l_2}(t)\right) \\ -\,\pi_{i_1}\left(\varepsilon_{i_1}(t)\right)\end{array}\right]. \tag{S.5}$$

Due to the strictly monotonically increasing of $\pi_i(s_i(t))$ with respect to $s_i(t)$, $\sum_{l_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)} s_{l_2}^*\pi_{l_2}\left(s_{l_2}^* - \varepsilon_{l_2}(t)\right) < \sum_{l_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)} s_{l_2}^*\pi_{l_2}\left(s_{l_2}^*\right) = \overline{\pi}(\boldsymbol{s}^*)$ holds. According to Definition 1, we have $\overline{\pi}(\boldsymbol{s}^*) \leq \pi_{i_1}(0) < \pi_{i_1}\left(\varepsilon_{i_1}(t)\right)$ since $\boldsymbol{s}^*$ corresponds to the Nash equilibrium as revealed by Theorem 3. Hence, the coefficient of the first-order term of $\varepsilon_{i_1}(t)$ is strictly negative, i.e., $A = \sum_{l_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)} s_{l_2}^*\pi_{l_2}\left(s_{l_2}^* - \varepsilon_{l_2}(t)\right) - \pi_{i_1}\left(\varepsilon_{i_1}(t)\right) < 0$. Hence, (S.5) has the solution form $\varepsilon_{i_1}(t) = \varepsilon_{i_1}(0)\exp\{\sigma A t\}$, which indicates that $\varepsilon_{i_1}(t) \to 0$ as $t \to \infty$.

ii) For any $i_2 \in \overline{\mathcal{A}}(\boldsymbol{s}^*)$, using the similar way above we can derive $ds_{i_2}(t)/dt = d\left(s_{i_2}^* - \varepsilon_{i_2}(t)\right)/dt = -d\varepsilon_{i_2}(t)/dt$. Thus, we further have

$$\frac{d\varepsilon_{i_2}(t)}{dt} = \sigma\left(\varepsilon_{i_2}(t) - s_{i_2}^*\right)\left[\overline{\pi}\left(\boldsymbol{s}(t)\right) - \pi_{i_2}\left(s_{i_2}^* - \varepsilon_{i_2}(t)\right)\right]. \tag{S.6}$$

Note that $\boldsymbol{s}^*$ is an ESS as shown in Theorem 4. It follows the definition of the ESS given in Definition 2 that $\overline{\pi}\left(\boldsymbol{s}(t)\right) > \overline{\pi}\left(\boldsymbol{s}^*\right)$ always holds for any point in its neighborhood $\boldsymbol{s}(t) \in \mathcal{N}(\epsilon)$ and $\boldsymbol{s}(t) \neq \boldsymbol{s}^*$. Additionally, the Nash equilibrium property also indicates $\overline{\pi}\left(\boldsymbol{s}^*\right) = \pi_{i_2}\left(s_{i_2}^*\right) > \pi_{i_2}\left(s_{i_2}^* - \varepsilon_{i_2}(t)\right)$ for all $i_2 \in \overline{\mathcal{A}}\left(\boldsymbol{s}^*\right)$. Therefore, we see $\overline{\pi}\left(\boldsymbol{s}(t)\right) - \pi_{i_2}\left(s_{i_2}^* - \varepsilon_{i_2}(t)\right) > 0$. Noting $\varepsilon_{i_2}(t) \in \left[0, s_{i_2}^*\right)$, the right-side term of (S.6) is always negative, i.e., $d\varepsilon_{i_2}(t)/dt < 0$. We formulate a function of $\varepsilon_{i_2}(t)$ as $V\left(\varepsilon_{i_2}(t)\right) = \left(\varepsilon_{i_2}(t)\right)^2$. It can be easily found that

- $V\left(\varepsilon_{i_2}(t)\right) = 0$ if and only if $\varepsilon_{i_2}(t) = 0$,
- $V\left(\varepsilon_{i_2}(t)\right) > 0$ if $\varepsilon_{i_2}(t) \in \left[0, s_{i_2}^*\right) \backslash \{0\}$, and
- $dV\left(\varepsilon_{i_2}(t)\right)/dt = 2\varepsilon_{i_2}(t)d\varepsilon_{i_2}(t)/dt < 0$ if $\varepsilon_{i_2}(t) \in \left[0, s_{i_2}^*\right) \backslash \{0\}$ while $dV\left(\varepsilon_{i_2}(t)\right)/dt = 0$ for $\varepsilon_{i_2}(t) = 0$.

As a consequence, $V\left(\varepsilon_{i_2}(t)\right)$ is a Lyapunov function candidate and $\varepsilon_{i_2}(t) = 0$ is asymptotically stable, i.e., from any initial point $\varepsilon_{i_2}(0) \in \left[0, s_{i_2}^*\right) \backslash \{0\}$, $\varepsilon_{i_2}(t) \to 0$ as $t \to \infty$.

Combining both the results above, we can conclude that the socially optimal solution $\boldsymbol{s}^*$ is asymptotically stable.

## 8 Proof of Lemma 1

For the sake of simplicity, we let $g_i(\boldsymbol{s}) = s_i \geq 0$ and $g_{i+M}(\boldsymbol{s}) = 1 - s_i \geq 0$ for $i = 1, 2, \ldots, M$ be the equivalent inequality constraints $\boldsymbol{s} \in [0,1]^M$ and denote by $\mathcal{A}(\boldsymbol{s})$ the active set at any feasible point $\boldsymbol{s}$ that consists of the indices of the inequality constraints $l$ for which $g_l(\boldsymbol{s}) = 0$,

i.e., $\mathcal{A}(\boldsymbol{s}) = \{l|g_l(\boldsymbol{s}) = 0, l = 1, 2, \ldots, 2M\}$. Besides, let the equality constraint of (9) be $h(\boldsymbol{s}) = 1 - \sum_{i=1}^{M} s_i$. According to **Theorem 1**, since the Nash equilibrium of the game model (9), $\boldsymbol{s}^{\mathrm{NE}}$, is the global minimizer of the convex optimization problem, it must satisfy the Karush-Kuhn-Tucker (KKT) conditions, i.e., the first-order necessary optimality conditions for a feasible solution to be optimal in a constrained optimization problem. That is, there always exist constants $w_l \in \mathbb{R}_{\geq 0}$ for $l \in \mathcal{A}(\boldsymbol{s}^{\mathrm{NE}})$ and $v \in \mathbb{R}$, such that the following equality is always held at $\boldsymbol{s}^{\mathrm{NE}}$

$$\bigtriangledown_{\boldsymbol{s}} F\left(\boldsymbol{s}^{\mathrm{NE}}\right) = \sum_{l \in \mathcal{A}(\boldsymbol{s}^{\mathrm{NE}})} w_l \bigtriangledown_{\boldsymbol{s}} g_l\left(\boldsymbol{s}^{\mathrm{NE}}\right) + v \bigtriangledown_{\boldsymbol{s}} h\left(\boldsymbol{s}^{\mathrm{NE}}\right). \tag{S.7}$$

On the other hand, the feasibility of any feasible point $\widetilde{\boldsymbol{s}}$ guarantees $g_l\left(\widetilde{\boldsymbol{s}}\right) \geq 0$ for all $l$. Because $g_l\left(\widetilde{\boldsymbol{s}}\right)$ is a linear function of $\widetilde{\boldsymbol{s}}$, applying the first-order Taylor's theorem to it leads to $g_l\left(\widetilde{\boldsymbol{s}}\right) = g_l\left(\boldsymbol{s}^{\mathrm{NE}}\right) + \left(\bigtriangledown_{\boldsymbol{s}} g_l\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right)$. Recall $g_l\left(\boldsymbol{s}^{\mathrm{NE}}\right) = 0$ for all $l \in \mathcal{A}\left(\boldsymbol{s}^{\mathrm{NE}}\right)$. We can get

$$g_l\left(\widetilde{\boldsymbol{s}}\right) = \left(\bigtriangledown_{\boldsymbol{s}} g_l\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right) \geq 0 \tag{S.8}$$

for all $l \in \mathcal{A}\left(\boldsymbol{s}^{\mathrm{NE}}\right)$. $h\left(\boldsymbol{s}^{\mathrm{NE}}\right)$ is also a linear function. Thus, we have $h\left(\widetilde{\boldsymbol{s}}\right) = h\left(\boldsymbol{s}^{\mathrm{NE}}\right) + \left(\bigtriangledown_{\boldsymbol{s}} h\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right)$. The equality constraints $h\left(\boldsymbol{s}^{\mathrm{NE}}\right) = 0$ and $h\left(\widetilde{\boldsymbol{s}}\right) = 0$ satisfied at the Nash equilibrium $\boldsymbol{s}^{\mathrm{NE}}$ and the feasible $\widetilde{\boldsymbol{s}}$ further result in

$$\left(\bigtriangledown_{\boldsymbol{s}} h_l\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right) = 0. \tag{S.9}$$

Combining (S.7) with (S.8) and (S.9) can get

$$\begin{aligned}
&\left(\bigtriangledown_{\boldsymbol{s}} F\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right) \\
&= \sum_{l \in \mathcal{A}(\boldsymbol{s}^{\mathrm{NE}})} w_l \left(\bigtriangledown_{\boldsymbol{s}} g_l\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right) \\
&\quad + v\left(\bigtriangledown_{\boldsymbol{s}} h\left(\boldsymbol{s}^{\mathrm{NE}}\right)\right)^{\mathrm{T}}\left(\widetilde{\boldsymbol{s}} - \boldsymbol{s}^{\mathrm{NE}}\right) \geq 0,
\end{aligned} \tag{S.10}$$

which is indeed the result of **Lemma 1**. Additionally, it is worth pointing out that the form of (S.10) can be viewed as the variational inequality characterization of the Nash equilibrium in the model (9).

## 9 Proof of Lemma 2

We substitute the expressions of $f_i(s_i, \boldsymbol{s})$ and $f_i(s_i', \boldsymbol{s}')$ into (13) to get the following inequality

$$\frac{s_i' s_i}{(1-\alpha_i)^{N s_i - 1}} \leq \lambda \frac{s_i'^2}{(1-\alpha_i)^{N s_i' - 1}} + \mu \frac{s_i^2}{(1-\alpha_i)^{N s_i - 1}}. \tag{S.11}$$

To prove (13) is then equivalently to prove (S.11). When $s_i' = 0$, it can be easily observed that (S.11) is satisfied with $\mu > 0$. In this case, **Lemma 2** is held. Otherwise, we examine the case of $s_i' \in (0, 1]$. To simplify the mathematical representation, let $\nu = s_i / s_i'$. With $\nu$, we can rearrange (S.11) as following

$$0 \leq \lambda(1-\alpha_i)^{N(s_i - s_i')} + \mu\nu^2 - \nu. \tag{S.12}$$

$(1-\alpha_i)^{N(s_i - s_i')} \geq (1-\alpha_i)^N$ is always held with $1 - \alpha_i \in (0, 1)$ and $s_i, s_i' \in [0, 1]$. Additionally, since $\mu \in (0, 1)$ as given in the condition of **Lemma 2**, $\mu\nu^2 - \nu$ is a convex function of $\nu$, whose minimizer is obviously $-\frac{1}{4\mu}$. In other words, $\mu\nu^2 - \nu \geq -\frac{1}{4\mu}$ is held with $\mu \in (0, 1)$. Combining the results aforementioned with the condition $\mu\lambda \geq \frac{1}{4}(1-\alpha_i)^{-N}$ indicates $\lambda(1-\alpha_i)^{N(s_i - s_i')} + \mu\nu^2 - \nu \geq \lambda(1-\alpha_i)^N - \frac{1}{4\mu} \geq \frac{1}{4\mu}(1-\alpha_i)^{-N}(1-\alpha_i)^N - \frac{1}{4\mu} = 0$. This means that (S.12) is held and **Lemma 2** is proven.

## 10 Proof of Lemma 3

According to the expression of $q_i\left(s_i, \boldsymbol{s}\right)$ as in (14), the channel price explicitly involves the variable $s_i$. Thus, it is obvious that $\left|\partial q_i\left(s_i, \boldsymbol{s}\right)/\partial s_l\right| = 0$ for any $l \neq i$ and $l \in \mathcal{M}$, while $\left|\partial q_i\left(s_i, \boldsymbol{s}\right)/\partial s_i\right|$ is finite due to the fact that $s_i$ is ranging within a closed interval $[0, 1]$.

## 11 Proof of Lemma 4

It follows the definition $W_k = k - W\left\lfloor \frac{k}{W} \right\rfloor$ that

$$\frac{W_k}{W} = \frac{k}{W} - \left\lfloor \frac{k}{W} \right\rfloor \leq 1. \tag{S.13}$$

(S.13) always holds according to the basic property of the floor function, which can further indicate $W_k \leq W$.

## 12 Proof of Lemma 5

According to the definition of $\Delta s_i(k)$, the result of **Lemma 5** holds when $\overline{\pi}^{W_k}(\boldsymbol{s}(k)) - \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k)) = 0$. Next, we need only investigate the two cases where $\overline{\pi}^{W_k}(\boldsymbol{s}(k))$ and $\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$ are not identical:

i) If $\overline{\pi}^{W_k}(\boldsymbol{s}(k)) > \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$, we can easily get

$$\begin{aligned}
&\left|\Delta s_i(k)\right| = \sigma s_i(k)\left[\overline{\pi}^{W_k}(\boldsymbol{s}(k)) - \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))\right] \\
&\leq \sigma s_i(k)\overline{\pi}^{W_k}(\boldsymbol{s}(k)) = s_i(k)\sum_{l=1}^{M} \sigma s_l(k)\pi_l^{W_k}\left(s_l(k), \boldsymbol{s}(k)\right) \\
&< s_i(k)\sum_{l=1}^{M} s_l(k) = s_i(k).
\end{aligned} \tag{S.14}$$

The last inequality of (S.14) follows the condition of $\sigma$ given in (18), i.e., $\sigma\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k)) < 1$ for all $i \in \mathcal{M}$, and $\sum_{l=1}^{M} s_l(k) = 1$.

ii) As for $\overline{\pi}^{W_k}(\boldsymbol{s}(k)) < \pi_i^{W_k}(s_i(k), \boldsymbol{s}(k))$, we can also have in the similar way above

$$\begin{aligned}
&\left|\Delta s_i(k)\right| = \sigma s_i(k)\left[\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k)) - \overline{\pi}^{W_k}(\boldsymbol{s}(k))\right] \\
&\leq \sigma s_i(k)\pi_i^{W_k}(s_i(k), \boldsymbol{s}(k)) < s_i(k).
\end{aligned} \tag{S.15}$$

Thus, **Lemma 5** is proven.

## 13 Proof of Lemma 6

Using the mean value theorem we can represent the difference of any two functions $\pi_i(s_i, \boldsymbol{s})$ and $\pi_i(\tilde{s}_i, \tilde{\boldsymbol{s}})$ with respect to any $\boldsymbol{s}, \tilde{\boldsymbol{s}} \in \Delta^{M-1}$, respectively, as $\pi_i(s_i, \boldsymbol{s}) - \pi_i(\tilde{s}_i, \tilde{\boldsymbol{s}}) = \left(\partial\pi_i(s_i, \boldsymbol{s})/\partial\boldsymbol{s}|_{\boldsymbol{s}=\boldsymbol{x}}\right)^{\mathrm{T}}(\boldsymbol{s} - \tilde{\boldsymbol{s}})$ with an existing real number $\rho \in (0, 1)$, where $x_i = (1-\rho)s_i + \rho\tilde{s}_i$ and $\boldsymbol{x} = (1-\rho)\boldsymbol{s} + \rho\tilde{\boldsymbol{s}}$. Based on this, we apply the Cauchy-Schwarz inequality to derive the upper bound of the difference as

$$\left|\pi_i(s_i, \boldsymbol{s}) - \pi_i(\tilde{s}_i, \tilde{\boldsymbol{s}})\right| \leq \left\| \left.\frac{\partial\pi_i(s_i, \boldsymbol{s})}{\partial\boldsymbol{s}}\right|_{\boldsymbol{s}=\boldsymbol{x}} \right\| \left\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\right\|. \tag{S.16}$$

Recalling **Lemma 3**, we have $\| \partial \pi_i(s_i, \boldsymbol{s}) / \partial \boldsymbol{s} |_{\boldsymbol{s}=\boldsymbol{x}} \| \leq \sqrt{M} Q$ and then derive

$$|\pi_i(s_i, \boldsymbol{s}) - \pi_i(\tilde{s}_i, \tilde{\boldsymbol{s}})| \leq \sqrt{M} Q \, \|\boldsymbol{s} - \tilde{\boldsymbol{s}}\| \,, \tag{S.17}$$

which coincides with the definition of the Lipschitz continuity.

## REFERENCES


[1] R. W. Rosenthal, "A class of games possessing pure-strategy nash equilibria," *International Journal of Game Theory*, vol. 2, no. 1, pp. 65–67, Dec 1973. [Online]. Available: https://doi.org/10.1007/BF01737559

[2] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bulletin of the American Mathematical Society*, vol. 40, no. 4, pp. 479–519, 2003.